PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C12N 15/12, C07K 14/705, 19/00, G01N 33/68, C12N 15/10

A1

(11) International Publication Number: WO 00/60070

(43) International Publication Date: 12 October 2000 (12.10.00)

(21) International Application Number: PCT/EP99/02283

(22) International Filing Date: 1 April 1999 (01.04.99)

(71) Applicant *(for all designated States except US)*: INNOGENETICS N.V. [BE/BE]; Industriepark Zwijnaarde 7, Box 4, B–9052 Ghent (BE).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: DESMET, Johan [BE/BE]; Lode de Boningelaan 12, B–8500 Kortrijk (BE). HUFTON, Simon [NL/NL]; Moselborg 5, NL–6228 BJ Maastricht (NL). HOOGENBOOM, Hendricus [NL/NL]; Hertogsingel 46, NL–6214 AE Maastricht (NL). SABLON, Erwin [BE/BE]; Robbroeklaan 1A, B–1785 Merchtem (BE).

(74) Common Representative: INNOGENETICS N.V.; Jacobs, Philippe, Industriepark Zwijnaarde 7, Box 4, B–9052 Ghent (BE).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) **Title:** A POLYPEPTIDE STRUCTURE FOR USE AS A SCAFFOLD

(57) **Abstract**

The present invention relates to a polypeptide chain having a β sandwich architecture which can be used as a scaffold, i.e. a supporting framework, carrying antigen– or receptor–binding fragments. More specifically, the present invention describes a non–glycosylated CTLA4–like β sandwhich carrying new and randomized peptide sequences.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

A POLYPEPTIDE STRUCTURE FOR USE AS A SCAFFOLD

FIELD OF THE INVENTION

The present invention relates to a polypeptide chain having a β sandwich architecture which can be used as a scaffold, i.e. a supporting framework, carrying antigen- or receptor-binding fragments. More specifically, the present invention relates to a β sandwich derived from the naturally occuring extracellular domain of CTLA-4.

BACKGROUND OF THE INVENTION

The primary amino acid sequence of a protein determines its three-dimensional (3D) structure, which in turn determines protein function. In most cases the 3D structure is essentially uneffected by amino acids at loci where the amino acid side chain is directed towards the solvent. Loci where limited variety is allowed have the side group directed toward other parts of the protein. This general rule, however, does not hold for those proteins whose function consists of, or relies upon binding with other biomolecules. Binding of such proteins to other molecules is non-covalently but often very tightly and specifically. In such instance, the identity of residues of the interacting domains that point towards the solvent becomes very important as well as the 3D-structure that they adopt, because binding results from complementarity of the surfaces that come into contact: bumps fit into holes, dipoles align and hydrophobic residues contact other hydrophobic residues. Individual water molecules eventually can fill a well determined space in intermolecular interfaces, and then usually form hydrogen bonds to one or more atoms of the protein or other biomolecules or to other bound water.

The 3D structure can often be predicted on the basis of the primary amino acid sequence. A major obstacle towards the correct prediction of the entire protein structure is the modelling of "loops". These are regions of the polypeptide chain that have a lower degree of structural and sequential regularity and where insertions and deletions often occur during evolution. Yet, this diversity does not fundamentally alter the core structure

of the remaining polypeptide chain. Often such loops, which are on the outside of the protein structure, will determine the interaction with other biomolecules. The conformation of loop fragments is largely controlled by the attachment of their ends to fragments that constitute the framework of the molecule, for instance secondary structural units. This is also the basic idea when grafting such polypeptide chains onto the core of the protein structure which functions as a scaffold. Scaffolds, which are able to restrain the 3D structure of a polypeptide chain within a loop, can thus be of utmost importance.

Immunoglobulins are multi domain proteins consisting of a heavy and a light chain. The variable domain of immunoglobulins is composed of an antigen binding site formed by six loops clustered in space and carried by two domains that adopt a so called-β sandwich structure (one β domain from the heavy chain carrying 3 loops and one from the light chain carrying 3 loops) (Amit et al., 1986). The high sequence variability of these loops allows immunoglobulins to recognise a variety of antigens. Though this sequence variation does determine in part the structure of the loops and the antigen specificity, the gross structure of the immunoglobulin variable domain's framework is hereby not altered. Reversely, the latter β-domains, consisting of anti-parallel strands that are connected by the loops, function as a scaffold that strongly determines the three dimensional structure of the loops by constraining their structure and providing a carrier function (Amzel,L.M. and Poljak,R.J. , 1979; Williams A.F.,1987). Furthermore, the β-sandwich architecture of both the heavy and light chains have a roughly half-cylindrical structure, with a more or less open side. In functional immunoglobulins, both domains (heavy and light chain) are packed onto eachother thereby forming a barrel-like fold.

A serious drawback in using the immunoglobulin scaffold to restrain random polypeptide sequences, however, is that the core β-sandwich is composed of two chains (the heavy and light chains) which interact, thereby bringing the variable loops of both chains (6 in total) in close proximity. This complicates the use of immunoglobulins for restraining the conformation of randomized polypeptide sequences such as in antibody libraries on filamentous phage. Large peptide libraries on phage are described in US 5223409 to Ladner et al. and US 5571698 to Ladner et al. In the latter system, each randomized peptide is fused to the gene III protein of the M13 phage. In this regard, it is

clear that a large, multi-domain protein is not ideal in phage-derived selection of peptidic leads. Consequently, efforts have been devoted to design novel, smaller and simpler proteins, based on the structure of the variable domain of an antibody. One of these efforts is the construction of a protein, named minibody, which consists of a fusion of the variable regions of an antibody and the CH3 domain of an immunoglobulin, all in a single polypeptide. Minibodies spontaneously dimerize to form an antibody fragment that is bivalent. The minibody scaffold, which contains a number of residue modifications with respect to the original immunoglobulin sequence, can be obtained both by solid-phase synthesis and by expression in bacteria (Bianchi et al, 1993; Bianchi et al ,1994). Further modifications of the latter molecule were successful in improving some important properties, for example its solubility in aqueous media (Bianchi et al ,1994). Other attempts to simplify the existing multi domain immunoglobulin molecule have resulted in recombinant fragments, made of two forms: Fv fragments and Fab's. Fv fragments are heterodimers composed of a variable heavy chain (V_H) and a variable light chain (V_L) domain and are the smallest functional fragments of antibodies that maintain the binding and specificity of the whole antibody. In these constructs, the immunoglobulin domains forming the characteristic β-barrel can still interact with each other. However, Fv fragments are unstable. Stable Fv's have been produced by making recombinant molecules in which the V_H and V_L domains are connected by a peptide linker so that the antigen-binding site is regenerated in a single molecule. These recombinant molecules are termed single chain Fv's (scFv's; Raag and Whitlow, 1995). The V_H-V_L heterodimer can alternatively be stabilized by an interchain disulfide bond and is then termed dsFvs (Reiter et al, 1994). Fab fragments, on the other hand, are composed of the light chain and the heavy chain Fd fragment (V_H and CH1) connected to each other via the interchain disulfide bond between CL and CH1.

Taken together, each simplification of the otherwise complex multidomain immunoglobulin molecule still results in a unit wherein two immunoglobulin folds interact with each other thereby reconstituting the antigen binding site which consists of two pairs of three loops each grafted onto a separate β-sandwich framework.

Another drawback for using immunoglobulins to restrain randomized polypeptide

sequences is the fact that the antigen-binding site of the latter molecules, which is constituted of six clustering loops, may not be able to bind small and/or cryptic sites within molecules. In other words, steric hindrance during the screening or selection for proteins that bind small and/or cryptic sites may occur when using an immunoglobulin framework. Solutions to the latter problem have been offered by restraining randomized polypeptide sequences of variable length by means of cyclization upon for instance incorporation of cysteine residues at fixed positions forming a disulfide bond, or, by means of incorporating a metal binding motif or a conserved hydrophobic core into the polypeptide, or, by means of embedding peptide segments in "small proteins" with a defined 3D structure such as endothelin, trefoil proteins, guanylin and the like (Ladner, 1995; Cannon et al. 1996). While these solutions have been very elegant, they result in small peptides which may rather quickly be cleared from the bloodstream, thus posing problems for certain therapeutic purposes.

The present invention relates to an alternative, simpler version of a monomeric polypeptide, compared to antibody-derived and small protein scaffolds, which can, to our surprise, efficiently be used as a scaffold.

AIMS OF THE INVENTION

The more tightly a polypeptide segment is constrained, the more likely this segment will bind to its target with high specificity and affinity (Ladner, 1995). In other words, there is a need to design scaffolds which are able to carry, or are embedded with, polypeptides efficiently binding to an antigen or receptor. Already existing scaffolds, such as antibody-derived scaffolds and "small proteins", are still faced with particular limitations. Indeed, antibody-derived scaffolds containing 6 loops constituting the antigen-binding site, are composed of rather large and complex dimers which, due to steric hindrance, may not be usefull for carrying polypeptides binding small and/or cryptic sites. Furthermore, large molecules containing an antigen binding site composed of six loops have a relatively higher chance to bind in a non-specific manner, or to cross-react, compared to smaller and simpler molecules containing an antigen binding site

composed of less than six loops (Maclennan, 1995). Moreover, antibody-derived scaffolds have the tendency to denature more easily during the aggressive purification protocols required for biotherapeutic production (Maclennan, 1995). "Small proteins", on the other hand, may be quickly cleared from the bloodstream which is not desirable for certain therapeutic purposes. Furthermore, both antibody-derived -and "small protein" scaffolds, which are used to generate large peptide libraries using phage-display technology, are limited by the fact that phage display is not able to display active forms of proteins that require eukaryotic-specific posttranslational modifications (such as glycosylation) for activity (Cannon et al., 1996). Also Metzler et al. (1997) indicated that glycosylation is important for binding activity, structural integrity and solubility of CTLA-4.

In order to overcome the above-indicated limitations, it is an aim of the present invention to provide an alternative, simplier and preferably unglycosylated scaffold which is not easily cleared from the blood stream.

It is also an aim of the present invention to provide a scaffold structure which is sufficiently stable to allow grafting of polypeptide chains as loops onto said scaffold without substantial alteration of the scaffold structure.

It is further an aim of the present invention to provide a scaffold structure that can restrain randomized polypeptide sequences thereby determining the 3D-structure of said polypeptide sequences.

It is also an aim of the present invention to provide a scaffold structure wherein at least one, preferably two and more preferably three of the restrained polypeptide loops are constituting a binding domain for other molecules.

It is also an aim of the present invention to provide a scaffold structure wherein at least two, preferably three, more preferably four, more preferably five and even more preferably six of the restrained polypeptide loops are constituting at least two binding domains for other molecules. This aim more particularly relates to providing a scaffold structure that allows for the design of bispecific molecules wherein both binding domains are located on the opposite sides of the molecule.

It is also an aim of the present invention to provide a scaffold structure which is

soluble under physiological conditions.

It is also an aim of the present invention to provide an unglycosylated scaffold structure that remains soluble and/or functional in a bacterial background. In this regard, it should be clear, in view of the prior art (i.e. see Cannon et al., 1996, and, Metzler et al., 1997) which indicates that unglycosylated molecules may aggregate and/or may have a reduced bioactivity, that a person skilled in the art is guided away to use an unglycosylated scaffold.

It is also an aim of the present invention to provide a scaffold structure that is essentially functional as a monomeric protein.

All these aims are met by the following embodiments of the present invention.

BRIEF DESCRIPTION OF DRAWINGS

Figure 1 shows a schematic representation of the scaffold of the present invention. The six basic β-strands are named S1, S2, S3, S4, S5 and S6; the three additional or optional β-strands are named A1, A2 and A3. The loops connecting the strands are indicated by a thin, black, curving line.

Figure 2 shows the sequence coding for SCA, which is based on the extracellular part of human CTLA-4 gene as described by Metzler *et al.* (1997) but without the C-terminal cysteine residue (Cys 123). β-sheets are underlined, Cysteine residues and glycosylation sites are shown. Mutations and there effect on CD80/86 binding are also indicated.

Figure 3 indicates the specific primers used to clone SCA as a *Apa*L1/*Not*1 fragment.

Figure 4 shows a shematic representation of the Phagemid pCES1 construct. Antibody genes : V_L-C_L, variable (V) and constant (C) region of the light chain; V_H-C_{H1}, variable and first constant region of the heavy chain; P*lacZ*, promoter; rbs, ribosome binding site; S, signal sequence; H6, six histidines stretch for IMAC purification; tag, *c-myc*-derived tag; amber, amber codon that allows production of soluble Fab fragments in non-

suppressor strains, gIII, gene encoding one of the minor coat proteins of filamentous phage. Restriction sites used for cloning are indicated.

Figure 5 demonstrates the expected patterns of the *Bst*N1 fingerprints of clones with the correct sized insert.

Figure 6 demonstrates a strong signal showing binding to both B7.1-Ig and B7.2-Ig and not to BSA or plastic. The latter signal also decreases proportionately with the amount of phage present.

Figure 7 shows an SDS-page analysis with DTT (lanes 1 and 3) or without DTT (lanes 2 and 4) followed by detection with anti-CTLA-4 (lanes 1 and 2) or with anti-myc (lanes 3 and 4). which demonstrates that part of the expressed 16 kD SCA protein is present in the medium. The 83 kD band is probably due to a fusion of the SCA protein with the gene 3 coat protein of the phage.

Figure 8 indicates the specific primers used for amplification (oligo1 and oligo 2) and reamplification (oligo 1 and oligo 3) of the SCA.

Figure 9 schematic representation of the cloning strategy as described in example 6.

Figure 10a Diagrammic presentation of selection of phage on αvβ3 integrin using a polyclonal phage ELISA (see also figure 10b for the enrichment factor of the several selection rounds).

Figure 11 Diagrammic presentation of phage ELISA of selected clones on αvβ3 integrin

Figure 12 Alignment of the wild type CTLA-4 CDR3 loop and border sequences with the peptides obtained after selection and showing the RGD sequence and its selected borders

Figure 13 Diagrammic presentation of soluble ELISA of selected RGD clones on $\alpha v\beta 3$ integrin

Figure 14 Diagrammic presentation of binding of anti CTLA-4 antibodies with the scaffold derivatives

Figure 15 Facs analysis of binding to HUVECS cells

Figure 16 Diagrammic presentation of phage ELISA of scaffold derivatives with altered sequences in loops A1/S1 and S2/A2

DETAILED DESCRIPTION OF THE INVENTION

The invention described herein draws on previously published work and pending patent applications. By way of example, such work consists of scientific papers, patents or pending patent applications. All these publications and applications, cited previously or below are hereby incorporated by reference.

The present invention relates to a scaffold composed of a single-chain polypeptide having the following structural properties:

- it contains at least two cysteine residues which form at least one disulphide bond, and
- it possesses less than 10% of alpha helical conformation, and
- it contains at least six β-strands (S1, S2, S3, S4, S5, S6) which:
  - are connected by amino acid loops of variable conformation and lenght according to the following topology: S1-S2-S3-S4-S5-S6, and wherein at least one of said loops is different for at least three contiguous amino acid residues by way of substitution, deletion or insertion when compared to the naturally occurring CTLA-4 molecule,

- form two β-sheets (one formed by S1/S4/S3 and one formed by S6/S5/S2 wherein the symbol "/" denotes the hydrogen bonding interactions between two spatially adjacent β strands) which are each characterized by an anti-parallel arrangement of said β strands and which are packed onto eachother so that they form a β sandwich architecture.

Furthermore, the present invention regards a scaffold as defined above which may contain a maximum of three additional β strands (A1, A2 and A3), and, which has the following topology: A1-S1-S2-A2-A3-S3-S4-S5-S6, and wherein at least one of said loops is different for at least three contiguous amino acid residues by way of substitution, deletion or insertion when compared to the naturally occurring CTLA-4 molecule, and, which possesses the following β-sheets: A1/S1/S4/S3 and S6/S5/S2/A2/A3.

The present invention also relates to a scaffold as defined above wherein :

-A1 is the amino acid sequence AQPAVVLA (SEQ ID 1) or any functionally equivalent derivative of said sequence,

-S1 is the amino acid sequence ASFPVEY (SEQ ID 2) or any functionally equivalent derivative of said sequence,

-S2 is the amino acid sequence EVRVTVLRQA (SEQ ID 3) or any functionally equivalent derivative of said sequence,

-A2 is the amino acid sequence QVTEVCAA (SEQ ID 4) or any functionally equivalent derivative of said sequence,

-A3 is the amino acid sequence TYMMGNELTFLDDS (SED ID 5) or any functionally equivalent derivative of said sequence,

-S3 is the amino acid sequence ICTGTSS (SEQ ID 6) or any functionally equivalent derivative of said sequence,

-S4 is the amino acid sequence QVNLTIQ (SEQ ID 7) or any functionally equivalent derivative of said sequence,

- S5 is the amino acid sequence GLYICKVE (SEQ ID 8) or any functionally equivalent derivative of said sequence,

- S6 is the amino acid sequence GIGNGTQIY (SEQ ID 9) or any functionally equivalent derivative of said sequence.

The applicability of a scaffold lies in the ability to introduce diversity without destroying the tertiary structure of the protein fold and the ability to recover binding molecules from a diverse repertoire. The latter can be achieved by phage display and affinity selection on the ligand of choice. Although the properties of a scaffold are largely determined by the nature of the application, a list of important properties is shown below:

-the scaffold should be a small, globular protein,

-it should be a single domain (thus easier to produce, purify and engineer into multivalent or multispecific reagents),

-it should fold correctly and the tertiary structure should not be perturbed by the introduction of diversity,

-it should be stable, preferably *in vivo* if that is of relevance for the chosen application,

-it should have permissive loops , patches and/or surfaces for introducing diversity at a number of chosen sites wherein the nature of the required binding surface depends on the application,

-it should have a large accessible binding surface which has the potential to be affinity matured,

-it should be engineerable to make monospecific/bispecific/trispecific or multispecific molecules -soluble and expressed well in e.g bacteria,

-it should allow fusion at the N-and/or C-terminus,

-it should should be preferably non-immunogenic and human if to be used therapeutically

-and it should resistant to proteolysis.

A number of novel scaffolds have already been described. Perhaps the one in the most advanced state of development is the Z domain (Nord et al., 1997). Others include, the minibody (Pessi et al., 1993; Martin et al. 1994), tendamistat (McConnell and Hoess., 1995) , zinc finger (Choo et al., 1995; Hamers-Casterman et al., 1993), cytochrome b_{562} (Ku and Schultz, 1995), trypsin inhibitor (Rottgen and Collins, 1995), synthetic coiled coil (Houston et al., 1996), conotoxins, thioredoxin (Colas et al., 1996), knottins (Smith et al., 1998), green fluorescent protein (Abedi et al., 1998), and fibronectin (Koide et al., 1998). For the generation of binding ligands there is probably no ideal scaffold for all

potential applications, the choice being governed by a changing number of the parameters cited above. It can be easily envisaged that the final application should have the most prominent role in this choice.

The disadvantages when antibodies are used as scaffolds are that:

-the binding surface is composed of two domains,

-the structure can not easily be inferred from the primary sequence,

-the construction of multimeric and multispecific reagents is complex and is limited by the two domain nature of an antibody molecule,

-the *in vivo* stability is not optimal,

-the performance of the antibody depends on the interactions between the light and heavy chains which cannot be predicted,

The disadvantages cited above are not encountered when the molecule of the present invention is used as a scaffold. Moreover, the molecule of the present invention has, in comparison to antibodies, the following advantages for use as a scaffold:

-it has monomeric nature which facilitates easier engineering of therapeutic reagents and allows engineering into a multivalent and/or multispecific molecule,

-it has improved stability and pharmacokinetics in comparison to recombinant antibody fragments due to the second non-Ig disulphide bridge,

-it has a binding domain consisting of at most three loops (which is half the amount of loops that constitute the binding domains of human antibodies), which enables it to bind to cryptic sites that are otherwise not accessible by antibodies. The latter "smaller" binding domain of the current invention also decreases the possibility of binding to other sites thereby decreasing the possibility of aspecific binding, and in addition it also helps in the design of small molecule mimetics, which is not the case for antibodies.

-it allows for replacement of loops on the other side of the molecule (the opposite site of the β-sandwich of where the CDR loops are located) allowing an easy design of bispecific monomeric molecules.

-Furthermore, analogous to antibodies, the basic design of the molecule of the present invention is of human origin (the scaffold of the present invention can be considered to be derived from the naturally occurring extracellular domain of the molecule CTLA-4),

so that therapeutic applications can be envisaged, reducing the risk of immunogenicity.
-The molecule of the present invention also has been shown to function as a Fc molecule which suggests it can be engineered into a dimeric molecule that can recruit effector function,

-Moreover, the molecule of the present invention is a very small molecule when compared to the bispecific molecules that have been described in the art and that have been designed to bind two different molecules or substances. This small size results in the advantage that if the molecule of the current invention is used to design a bispecific molecule that the two sites to which this molecule will bind will be brought into a very close proximity which has advantages in for instance drug targetting.

It should be clear that the scaffold of the present invention can be prepared by any method known in the art such as classical chemical synthesis, as described by Houbenweyl (1974) and Atherton & Shepard (1989), or by means of recombinant DNA techniques as described by Maniatis et al. (1982).

The term "alpha helical conformation" refers to the conformation of a polypeptide segment where the main chain adopts a fragment of a right-handed helix (phi- and psi-angles both near 60°), wherein all of the C=O and N-H groups of the peptide linkages lay roughly parallel to the axis of the helix, each carbonyl group being hydrogen bonded to the fourth N-H group on up the chain. The number of amino acid per turn of the helix is 3.61. The pitch (repeat distance) of the helix is 0.541 nm.

The term "beta (β) strand" refers to the conformation of a polypeptide segment where the main chain adopts an extended conformation (phi- and psi-angles around -120° and +120°, respectively) leading to a relatively linear structure. A beta strand can be interrupted by one or two non-extended residues (for instance a beta bulge), while still be considered as a single strand, as long as the principal axes of the two consecutive sub-fragments do not form an angle greater than 90°.

The term "beta (β) sandwich architecture" refers to two beta-sheets that pack against eachother and in which the strands of both sheets form a so-called relative 'twist angle' of around 30° (within the range 20-50°).

The terms "any functionally equivalent derivative of said sequences" refer to any

variant or fragment of the amino acid (aa) sequences represented by SEQ ID 1 to 9 which does not result in a significant alteration of the global structural and functional properties of the scaffold of the present invention. The latter terms include, post-translational modifications of the aa sequences represented by SEQ ID 1 to 9 such as glycosylation, acetylation, phosphorylation, modifications with fatty acids and the like. Included within the definition are, for example, amino acid sequences containing one or more analogues of an aa (including unnatural aa's), amino acid sequences with substituted linkages, mutated versions of the amino acid sequences, peptides containing disulfide bonds between cysteine residues, biotinylated amino acid sequences as well as other modifications known in the art. Included within the definition are also those single amino acid substitutions which can further improve the solubility of the unglycosylated protein. Examples of equivalent derivatives, which are not intended to limit the scope of the present invention but are purely illustrative, are:

-for A1, the amino acid sequences TQPSVVLA (SEQ ID 10), TQPPVVLA (SEQ ID 11), SQPAVVLA (SEQ ID 12), KQSPLLVVD (SEQ ID 13), KQSPMLVVN (SEQ ID 14), KQSPMLVAY (SEQ ID 15), KQLPRLVVY (SEQ ID 16), AQRPLLIVA (SEQ ID 17), TQSPAIMSA (SEQ ID 18), QESGPGLV (SEQ ID 19) and the like,

-for S1, the amino acid sequences ASFPCEY (SEQ ID 20), ASFSCEY (SEQ ID 21), VSLSCRY (SEQ ID 22), VNLSCKY (SEQ ID 23), ATLVCNY (SEQ ID 24, VTMTCSA (SEQ ID 25), VTMTCKS (SEQ ID 26), VTITCKA (SEQ ID 27), VTMSCKS (SEQ ID 28), MKLSCVA (SEQ ID 29), LRLSCAT (SEQ ID 30), LKLSCAA (SEQ ID 31), RKLSCAA (SEQ ID 32), LSITCTV (SEQ ID 33), VKLSCTA (SEQ ID 34), VKISCKA (SEQ ID 35), LSITCTV (SEQ ID 36), VKISCKA (SEQ ID 37), VQISCKA (SEQ ID 38) and the like,

-for S2, the amino acid sequences EVRVTVLRQT (SEQ ID 39), EVRVTVLREA (SEQ ID 40), EFRASLYKGV (SEQ ID 41), EFRASLYKGA (SEQ ID 42), EFRASLHKGL (SEQ ID 43), EFRASLHKGT (SEQ ID 44), EVRMYMYQQK (SEQ ID 45), KNYLTWYQQK (SEQ ID 46), EVRVVWYQQK (SEQ ID 47), NFRLAWYQQK (SEQ ID 48), KNFLAWYQQK (SEQ ID 49), EVRMSWVRQS (SEQ ID 50), EVRMEWVRQP (SEQ ID 51), EVRMSWVRHT (SEQ ID 52), EVRMHWVRQA (SEQ

ID 53), EVRVNWVRQP (SEQ ID 54), EVRMNWVKQR (SEQ ID 55), EVRIEWVKQR (SEQ ID 56), EVRVHWVRQS (SEQ ID 57), EVRIEWVKER (SEQ ID 58) and the like,

-for A2, the amino acid sequences QVTEVCAT (SEQ ID 59), QMTEVCAT (SEQ ID 60), QVTEVCAG (SEQ ID 61), QMTEVCAM (SEQ ID 62), SDVEVCVG (SEQ ID 63), SAVEVCVV (SEQ ID 64), SAVEVCAV (SEQ ID 65), SAVEVCFI (SEQ ID 66), SPRLLIYD (SEQ ID 67), SPKLLIYW (SEQ ID 68), PPKLLIYG (SEQ ID 69), GLEWVAEI (SEQ ID 70), RLEWIAAS (SEQ ID 71), RLEWVATI (SEQ ID 72), GLEWVAYI (SEQ ID 73), GLEWLGMI (SEQ ID 74), GLEWIGRI (SEQ ID 75), GLEWIGWI (SEQ ID 76), GLEWLGMI (SEQ ID 77), GLEWIGEI (SEQ ID 78), GLEWIGEI (SEQ ID 79) and the like,

-for A3, the amino acid sequences TFTVKNTLGFLDDP (SEQ ID 80), TFTEKNTVGFLDYP (SEQ ID 81), TYMVEDELTFLDDS (SEQ ID 82), TYTVENELTFIDDS (SEQ ID 83), NGNFTYQPQFRSNAEF (SEQ ID 84), NGNFTYQPQFRPNVGF (SEQ ID 85), NGNFSHPHQFHSTTGF (SEQ ID 86), NGNHSHPLQSHTNKEF (SEQ ID 87), YGNYSQQLQVYSKTGF (SEQ ID 88), NGNYSHQPQFYSSTGF (SEQ ID 89), SWNMTHKINSNSNKEF (SEQ ID 90), TSNLASGVPV (SEQ ID 91), ASTRESGVPD (SEQ ID 92), ASTRHIGVPD (SEQ ID 93), RLNSDNFATHYAESVKG (SEQ ID 94), RNKGNKYTTEYSASVKG (SEQ ID 95), SNGGGYTYYQDSVKG (SEQ ID 96), SSGSSTLHYADTVKG (SEQ ID 97), WGDGNTDYNSALKS (SEQ ID 98), DPANGNIQYDPKFRG (SEQ ID 99), YPGSGNTKYNEKFKG (SEQ ID 100), LPGSGSTNYNEKFKG (SEQ ID 101), WGGGSIEYNPALKS (SEQ ID 102), LPGSGRTNYREKFKG (SEQ ID 103) and the like,

-for S3, the amino acid sequences FCSGTFN (SEQ ID 104), TCIGTSR (SEQ ID 105), TCTGISH (SEQ ID 106), NCDGDFD (SEQ ID 107), NCDGNFD (SEQ ID 108), NCDGKLG (SEQ ID 109), NCTVKVG (SEQ ID 110), NCDGKLG (SEQ ID 111), DCDGKLG (SEQ ID 112), NCRGIHD (SEQ ID 113), RFSGSGS (SEQ ID 114), RFTGSGS (SEQ ID 115), RFAGSGS (SEQ ID 116), KFIISRD (SEQ ID 117), RFIVSRD (SEQ ID 118), RFTISRD (SEQ ID 119), RFTISRD (SEQ ID 120), RLSISFD (SEQ ID 121), KATITAD (SEQ ID 122), KATLTVD (SEQ ID 123), RLSISKD (SEQ ID 124),

KATFTAD (SEQ ID 125), and the like,

-for S4, the amino acid sequences RVNLTIQ (SEQ ID 126), KVNLTIQ (SEQ ID 127), TVTFRLW (SEQ ID 128), TVTFYLK (SEQ ID 129), TVTFYLQ (SEQ ID 130), SVTFYLQ (SEQ ID 131), TVTFYLR (SEQ ID 132), KVIFNLW (SEQ ID 133), SYSLTIS (SEQ ID 134), DFTLSIS (SEQ ID 135), DYTLTIS (SEQ ID 136), DFTLTIS (SEQ ID 137), RLYLQMN (SEQ ID 138), ILYLQMN (SEQ ID 139), TLFLEMT (SEQ ID 140), TLFLQMT (SEQ ID 141), QVFLKMN (SEQ ID 142), TAYLQL (SEQ ID 143), TAYMQLS (SEQ ID 144), QVFLKMN (SEQ ID 145), TATMQLS (SEQ ID 146), QIFLKMN (SEQ ID 147) and the like,

-for S5, the amino acid sequences GLYFCKVE (SEQ ID 148), GLYLCKVE (SEQ ID 149), GLYVCKVE (SEQ ID 150), DIYFCKIE (SEQ ID 151), DIYFCLKE (SEQ ID 152), ATYYCQQW (SEQ ID 153), AVYYCQNN (SEQ ID 154), ALYYCQQH (SEQ ID 155), AVYVCQND (SEQ ID 156), GIYYCVLR (SEQ ID 157), AIYYCARN (SEQ ID 158), GLYYCARR (SEQ ID 159), GMYYCARW (SEQ ID 160), ARYYCARE (SEQ ID 161), AVYYCATK (SEQ ID 162), AVYFCARG (SEQ ID 163), AVYYCARH (SEQ ID 164), AXYYCVSY (SEQ ID 165), AVYVCTRG (SEQ ID 166) and the like,

-for S6, the amino acid sequence GMGNGTQIY (SEQ ID 167), ERSNGTIIH (SEQ ID 168), EKSNGTIIH (SEQ ID 169), EKSNGTVIH (SEQ ID 170), TFGVGTKLE (SEQ ID 171), TFGAGTKLE (SEQ ID 172), TFGGGTKLE (SEQ ID 173), YWGQGTSVT (SEQ ID 174), VWGAGTTVT (SEQ ID 175), YWGRGTLVT (SEQ ID 176), YWGRGTLVT (SEQ ID 177), YWGQGTTLT (SEQ ID 178), YWGQGTTLT (SEQ ID 179), YWGQGTLVT (SEQ ID 180) and the like.

It should be clear that the scaffold of the present invention contains at least six β-strands (S1 to S6) and may, but does not have to, comprise one (A1 or A2 or A3), preferably two (A1 and A2, or, A1 and A3, or A2 and A3), and most preferably three (A1 and A2 and A3) additional β-strands (A1 to A3).

The terms "amino acid loops of variable conformation and length" refer to any aa sequence of any lenght and any conformation which is able to connect the β-strands of the current invention. The latter aa loops preferably contain at least one aa fragment which binds to a receptor or antigen (see further). In this regard, it should be clear that

not all loops need to contain (a) receptor/antigen binding fragment (s). Indeed, one out of the 6 loops may contain an above-indicated fragment; preferably, two out of the 6 loops contain an above-indicated fragment and more preferably, three out of the 6 loops contain an above-indicated fragment. However, also four, five and six out of the 6 loops may contain an above-indicated fragment. In the latter regard, the scaffold of the present invention may contain sets of fragments which are able to bind to more than one, preferably two, different receptors/antigens. For example, the loops which connect S5 and S6, S1 and S2, S3 and S4, and, A2 and A3 contain fragments binding to a tumor antigen or B7.1/B7.2 expressing cells and the loops connecting A3 and S3, S2 and A2, S4 and S5, and, A1 and B contain fragments binding to a toxin able to kill the tumor cell expressing said tumor antigen or said B7.1/B7.2 expressing cells in a manner similar as the one described in WO91/07437 to Pfreundschuh and WO96/40260 to De Boer & De Gast, respectively. Other specific examples of such amino acid loops, which are not intended to limit the scope of the present invention but are purely illustrative, are:

- loops connecting said β-strands A1 and S1: SSHGV (SEQ ID 181), SSRGI (SEQ ID 182), SSRGV (SEQ ID 183), SSRGV (SEQ ID 184), SNE (SEQ ID 185), NNE (SEQ ID 186), DNE (SEQ ID 187), DNA (SEQ ID 188), SPGEK (SEQ ID 189), QPGGS (SEQ ID 190) and the like,
- loops connecting said β-strands S1 and S2: SPSHNTD (SEQ ID 191), ASSHNTD (SEQ ID 192), ASPGKAT (SEQ ID 193), ESSGKAD (SEQ ID 194), ASHGKAT (SEQ ID 195), SYNLLAK (SEQ ID 196), TYNLFSK (SEQ ID 197), SYNLFSR (SEQ ID 198), TYNGTGK (SEQ ID 199), and the like,
- loops connecting said β-strands S2 and A2: ND (SEQ ID 200), DS (SEQ ID 201), GS (SEQ ID 202), NS (SEQ ID 203), PGS (SEQ ID 204), PGQ (SEQ ID 205), PEK (SEQ ID 206), PGK (SEQ ID 207), PDK (SEQ ID 208), and the like,
- loops connecting said β-strands S3 and S4: ES (SEQ ID 209), GN (SEQ ID 210), NE (SEQ ID 211), KD (SEQ ID 212), GT (SEQ ID 213), DSKS (SEQ ID 214), TSQS (SEQ ID 215), NAKN (SEQ ID 216), NPKN (SEQ ID 217), and the like,
- loops connecting said β-strands S4 and S5: GLRAVDT (SEQ ID 218), GLRAADT (SEQ ID 219), GLRAMDT (SEQ ID 220), GLSAMDT (SEQ ID 221), NLHVNHT (SEQ

ID 222), NLDVNHT (SEQ ID 223), NLYVNQT (SEQ ID 224), DLYVNQT (SEQ ID 225), NLFVNQT (SEQ ID 226), NMSASQT (SEQ ID 227), and the like,

- loops connecting said β-strands S5 and S6: LMYPPPYFV (SEQ ID 228), LMYPPPYYL (SEQ ID 229), LMYPPPYYV (SEQ ID 230), FMYPPPYLDN (SEQ ID 231), VMYPPPYLDN (SEQ ID 232), VLYPPPYIDN (SEQ ID 233), VMYPPPYIGN (SEQ ID 234), AMYPPPYVYN (SEQ ID 235), and the like.

The terms "homologous" and "homology" are used in the current invention as synonyms for "identical" and "identity"; this means that the amino acid sequences which are e.g. said to be 95 % homologous show 95 % identical amino acids in the same position upon alignment of the sequences.

It has to be understood that it is not an aim of the present invention to provide CTLA-4 molecules as such, nor molecules derived from CTLA-4 by means of single amino acid substitutions with the intention to increase solubility or alter the binding affinity and/or binding specificity of a CTLA-4 molecule for B7-molecules, the latter being present on the outer surface of antigen presenting cells. The structure as set out above therefore should have less then 98% homology with the amino acid sequence of the naturally occurring CTLA-4-molecule, preferably less then 97%, more preferably less then 95% and even more preferably less then 90% homology with the amino acid sequence of the naturally occurring CTLA-4-molecule. This is due to alterations within the loops which connect the several β-strands of the scaffold as set out above, and does not result from amino acid substitutions within the β-strands which could increase solubility or stability of the scaffold molecule without substantially altering the concept of the invention.

It is a preferred embodiment of the present invention that the amino acid sequence of at least one of the loops is different as compared to the amino acid sequence of the naturally occurring CTLA-4 molecule, preferably two loops and even more preferably three loops. A loop is considered different if at least three, more preferably four, and even more preferably five contiguous amino acid residues of the loop are absent (deletion), additionally present (insertion), or altered (substitution) when compared to the naturally occurring CTLA-4 molecule. Preferably the loop which is different as compared to the

naturally occurring CTLA-4 molecule is one of the loops which is not CDR1 nor CDR2 nor CDR3 when one would take the naturally occurring CTLA-4 molecule as a reference. This preferred embodiment therefore refers to the replacement of at least one of the loops which are connecting the A1 and S1, or S2 and A2, or A3 and S3, or S4 and S5 β-strands, or which are connecting the A1 and S1, or S2 and S3, or S4 and S5 β-strands of the scaffold if the additional β-strands A2 and A3 are ommitted. Replacement of these loops with randomized peptide sequences can not be considered to be analogous to the replacement of loops in Ig-like proteins as described in the art. The molecule of the present invention is not only different, moreover, the loops being referred to are located on the other side of the molecule as the CDR1, CDR2 and CDR3 loops which can be related to the antigen determining region of Ig-like proteins.

In an even more preferred embodiment the loop which is different as compared to the naturally occurring CTLA-4 molecule is one of the loops which is CDR1 or CDR2 or CDR3 when one would take CTLA-4 as a reference. This preferred embodiment more precisely refers to the replacement of at least one of the loops which are connecting the S1 and S2, or A2 and A3, or S3 and S4, or S5 and S6 β-strands, or which are connecting the S1 and S2, or S3 and S4, or S5 and S6 β-strands of the scaffold if the additional β-strands A2 and A3 are ommitted. Such molecules therefore do not display binding affinity altogether for B7 molecules present on the surface of antigen presenting cells.

The present invention further relates to a scaffold as defined above wherein said amino acid loops comprise fragments binding to a receptor or antigen.

The expression "fragments binding to a receptor or antigen" refers to any possible aa sequence which is part of the said loops (i.e. said fragments comprise maximally as many aa's as the loop itself and comprise, more frequently, less aa's than the loop wherefrom they are derived) and which binds any receptor or antigen. In this regard, it should be clear that the scaffold of the present invention can carry any randomized aa sequence.

The present invention concerns a single-chain polypeptide as defined above for use as a scaffold. In other words, the scaffold of the present invention can be used to

generate large peptide libraries which can be screened for peptides with desired binding characteristics as described in US 5223409 to Ladner et al. and US 5571698 to Ladner et al. In the two latter patents, a phage based system is used in which each randomized peptide is fused to the gene III protein of the M13 phage. However, also non-phage based systems such as screening peptides on polysomes or on the surface of *E. coli* (Tuerk & Gold, 1990, Science 249:505-510), or, using a system wherein each randomized peptide is fused to a DNA binding protein as described in US 5498530 to Schatz et al. can be used.

The scaffold of the present invention carrying carrying peptides selected from randomized aa sequences can be used for therapeutic, diagnostic, and related purposes. The man skilled in the art will appreciate that the targets or ligands bound by the recombinant proteins of the present invention can be proteins, nucleic acids, lipids and carbohydrates, or combinations thereof. It has to be understood that 'combinations thereof ' refers to glycoproteins, lipoproteins, and the like but also refers to dimeric (homo- or heterodimeric) biomolecules that are bound to each other by non-covalent means. Therefore, the term 'antigen' as used throughout the specification, has to be interpreted in the broadest sense. Because one can use an artificial screening system, the man skilled in the art will also appreciate that recombinant proteins can also be selected or screened for binding to xenobiotic molecules, that would otherwise be highly toxic for a biological system.

It has to be understood that the invention also relates to those recombinant proteins that are able to bind, as already indicated above, two entirely unrelated biomolecules or substances, by means of two binding surfaces, e.g. one surface consisting of the loops connecting the β-strands on one side of the β-sandwich and comprising the CDR loops in the natural protein, and one surface consisting of the loops that connect the β-strands on the other side of the β-sandwich. As such, both binding surfaces can operate independent from each other and selection or screening for such proteins can be achieved in two separate rounds:

- in the first round, the biomolecule that binds to one partner is retrieved from a library wherein the amino acid sequence of one set of loops has been randomized while the other

set of loops is untouched e.g. as in the naturally occurring protein,

- in the second round, the biomolecule that binds to a second partner is retrieved from a library wherein the amino acid sequence of the other set of loops has been randomized. In the second round, the randomized library may already contain those loops that bind for the first partner or, alternatively, remain untouched as in the naturally occurring protein. In the latter case, both sets of binding loops as retrieved which bind separate partners have to be recombined within a single scaffolding protein.The known and potential advantages of bispecific molecules is well known in the art. They provide key tools in therapeutic and diagnostic procedures such as immunohistochemistry (see example 11) and enzyme immunoassays, radiotherapy and immunotherapy, drug targetting, and for redirecting biomolecules or certain cell types to new sites. One binding partner can be used to target for certain cell types such as cancer cells inducing intracellular uptake, while the other binding partner can be a component of a life sustaining cellular mechanism of which the function is inhibited upon binding. One of the binding sites can also be a cytotoxic molecule, thereby forcing uptake of the cytotoxic molecule into the targeted cells. This strategy can also be used to force uptake of endogeneous molecules, such as auto-immune antibodies, toward cell types, such as liver cells. One binding partner can be used to target for certain cell types, while the other binding partner can be used to target for other cell types thereby forcing both cell types to stay in close proximity, such as any lymphocytes with for instance lymphocytes of a different type or with for instance cancer cells. For many aspects the molecule of the current invention has the additional advantage that it is a much smaller molecule then the bispecific antibodies already described in the prior art. This results in a much closer contact between for instance cancer cells and a radioactive isotope, or between one cell type and another, as for instance will be the case if cytotoxic T lymphocytes are redirected towards other cells by the molecule of the present invention. In such case the molecule of the present invention has been provided with a binding domain specific for a site on the targeted cells and a second CD3 domain or a CD3 like domain that recruits and activates the cytotoxic T lymphocytes.

In another preferred embodiment one of the binding sites of the scaffold itself can

also be used to function as a spacer or anchor in cases wherein immobilisation of the entire protein is desired, such as on membranes or on any carrier for chromatographic purposes, thereby allowing for a directed immobilisation. One of the loops can for instance comprise a repeat of Histidine residues or a His-tag thereby allowing anchorage or immobilisation to a carrier or membrane by means of a metal-chelate or by means of covalent linkage as for instance is described in patent application PCT/EP 98/03883. The man skilled in the art will understand that it is not a prerequisite that the two binding domains bind to different molecules. Both binding domains might as well bind identical molecules, thereby providing for instance a device for use in agglutination tests. One of the binding domains can be selected for recognizing red blood cells, while the other domain can be selected for binding to disease-specific antibodies, thereby providing a device for agglutination tests. It has to be understood that whenever the term 'His-tag' is used throughout the specification, a broad definition can be given to it, meaning any sequence that has an amount of histidine residues that is sufficiently high to allow appropriate application. In principle a 10 amino acids long sequence that has two histidine residues, and is repeated at least twice should suffice. More typical, said 'His-tag' contains at least three histidines within a stretch of 6 amino acids. The term 'His-tag' can also include any repetition of a histidine immediately followed by at least one residue. Preferably the term His-tag refers to at least three consecutive histidine residues, more preferably to at least four histidine residues, or even more preferably to at least five or at least six consecutive histidine residues.

In another preferred embodiment the present invention relates to the introduction of an RGD sequence, embedded within a randomized peptide sequence, within one of the loops of the scaffold, thereby making up a peptide library of RGD displaying scaffold molecules that can be selected for binding to integrin. Integrins are involved in a number of pathological processes that involve angiogenesis e.g thrombosis, osteoporosis, rheumatoid artheritus, diabetic retinopathy, cancer and atherosclerosis. $\alpha v\beta 3$ integrins are unique markers for activated angiogenic endothelium. As such they are ideal targets for antibody-based endothelial cell destruction. Specific targeting of the integrins on the tumor endothelial cell surface will have application as both imaging reagents, as

antitumor reagents and targets for gene therapy vectors which can readily access the tumor via the tumor vasculature. A number of peptides which can specifically target the tumor endothelium have been discovered (for example RGD peptides which bind to the αvβ3 integrins). Previous studies have clearly shown that peptides containing the RGD motif can specifically target αvβ3 integrins (Pasqualini et al., 1995; Pasqualini et al., 1997). The specificity, affinity and pharmacokinetics of these peptide ligands is sub-optimal and would greatly benefit from being constrained and presented in a larger scaffold. High affinity binding should require the positioning of the RGD motif at the apex of an extended loop with flanking residues which conformationally optimise the display of the motif. The RGD binding motif has previously been introduced into the CDR3 loop of an antibody scaffold with demonstrated high affinity binding (10pM) for αvβ3 integrin (Barbas et al., 1993)(Smith et al., 1994). Previous studies have also engineered Hirudisins as thrombin inhibitors with disintegrin activity by introducing an RGD sequence at the tip of the hirudins finger like structure. The antiplatelet activity that was observed was due to the introduced integrin directed RGD motif (Knapp et al., 1992). Due to the ubiquitous nature of integrins and related molecules, the key to the therapeutic potential of targeting integrins is to target them specifically. By presenting the RGD binding motif as a constrained specificity as part of a larger binder surface will allow the maturation of these molecules to have fine specificity to for example αvβ3 integrin. By sequentially introducing diversity and selecting for binding using phage display for example such binding molecules may be further matured for affinity and subtype specificity.

The present invention also regards, as stated above, a scaffold as defined above wherein said single-chain polypeptide is unglycosylated. However, it has to be understood that the scaffold of the present invention can also be glycosylated. Though one of the major advantages of the claimed scaffolds is the fact that they remain functional without glycosylation, allowing expression in bacterial systems, the fact of being unglycosylated is not a prerequisite for falling under the protected matter. Once a recombinant scaffold has been selected in a bacterial background for binding a specific molecule and thus as an unglycosylated protein, one can decide to glycosylate the scaffold

by expression in a eukaryotic background. This can have distinct advantages over the non-glycosylated scaffold for therapeutic purposes.

The present invention will now be illustrated by reference to the following examples which prove the concept of the invention and set forth particularly advantageous embodiments. However, it should be noted that these embodiments are illustrative and can not be construed as to restrict the invention in any way.

EXAMPLES

Example 1: Cloning of a recombinant scaffold, named SCA, for display on the surface of filamentous phage

The sequence coding for SCA, which is based on the extracellular part of human CTLA-4 gene as described by Metzler *et al.* (1997) but without the C-terminal cysteine residue (Cys 123), is cloned into the phage display vector pCES-1 (figure 4). The cysteine residue is not included because it may present some problems with the correct folding and presentation of the SCA- p3 fusion on filamentous phage. This cloning creates a translational fusion of SCA to the N-terminus of p3 for display on filamentous phage. SCA is cloned as a *Apa*L1/*Not*1 fragment (neither of these restriction sites are present in human CTLA-4) using specific primers (figure 3). Human CTLA-4 is amplified from 0.1ng vector pBSK(+)hCTLA-4 using primers CTLA-4 (Front) and CTLA-4 (Back) at 10pM, 10µl PCR buffer plus magnesium at 2.5mM, 5µl of dNTP's and 1µl of taq polymerase (Boerhinger Manheim EXPAND™). Amplification is done with a hot start of 10 minutes at 94°C and then for 30 cycles 1min 94°C, 1 min 50°C and 2 min 72°C. The correct sized PCR product is then purified (402bp) and digested with both *Apa*L1 and *Not*1. Vector pCES1 is also digested with *Apa*L1 and *Not*1. Both digested vector and digested SCA- insert are purified. Ligation is performed overnight at 16°C and is transformed into TG1. The *Bst*N1 fingerprints of clones with the correct sized insert, as assessed by PCR, demonstrate the expected patterns (Figure 5).

Example 2: Display of a recombinant, non-glycosylated and monomeric SCA-protein on the surface of filamentous phage

Phage displaying non-glycosylated SCA proteins on their surface are prepared as follows. Clones 2, 4 and 5 from figure 4 are grown to an OD of 0.5 in the presence of ampicillin at 100µg/ml and 2% glucose. Five mls of culture is infected with helper phage M13KO7 at a multiplicity of infection of 20:1 and incubated at 37°C for 30 minutes. Cells are resuspended in TY medium plus Amplicillin and Kanamycin (25µg/ml) and grown overnight at 30°C. Phages are recovered by precipitation by PEG and resuspended in PBS. Because the SCA protein is derived from the human CTLA-4 protein wherein the original CDR loops are retained, a binding is expected with B7.1 and B7.2 proteins. To demonstrate that such binding occurs, phages are titrated and a serial dilution is made to test for binding in ELISA to B7.1-Ig and B7.2-Ig. A strong signal is obtained showing binding to both B7.1-Ig and B7.2-Ig and not to BSA or plastic (Figure 6). The latter signal decreases proportionately with the amount of phage present. This demonstrates that un-glycosylated SCA has the proper conformation and remains functional as a monomer. There is no difference in binding behavior for B7-1 or B7-2 of SCA displaying phage.

Example 3: Soluble SCA - His6/myc expressed in the non-suppressor strain HB2151

The gene coding for SCA, preceded by the pelB signal sequence, is cloned into the phage display vector pCES-1 (figure 4) under control of the IPTG inducible lac-promoter. The coding sequence is then fused C-terminally to a His6 followed with a myc-tag. An amber stopcodon is located at the junction between this coding sequence and the sequence coding for the phage coat protein g3p. This allows one to choose whether the protein will be exposed on the surface of the viral coat through fusion with the g3p coat protein, or expressed as a soluble protein , without the need for a supplementary cloning step. When a suppressor type of strain is transformed with this construct, translation will continue through the stopcodon resulting in a fused SCA-g3p protein exposed on the surface of a tip of the phage. In a non-suppressor strain the amber stopcodon is recognized and translation stops. Subsequently, the His6- or myc-tagged

protein is transported through the cell membrane and accumulates in the periplasm. Non-suppressor cells HB2151 are transformed with the His6/myc- tagged construct and grown in liquid LB medium with ampicilline for selection and 1% glucose for metabolic repression of the lac-promoter at 28 °C until saturation is reached. This saturated culture is subsequently diluted 20 times in 20 ml LB+A+G and grown at 28 °C until an OD600 is reached of 0.5 to 0.6, which takes about three hours. The glucose containing medium is then removed and replaced by LB+A and 0.1 mM IPTG. The induced culture is further incubated for about 20 hours. The next day the OD600 is measured and the cells are fractionated.

Example 4: Fractionation of transformed cells

To examine whether the expressed SCA protein is actually secreted and accumulated in the periplasm, and possibly leaking out into the medium, a periplasmic fraction is prepared using a modified protocol as according to the osmotic shock procedure described by Neu and Heppel (1965).

Sixteen ml of cell culture is centrifuged en resuspended in 1 ml of icecold TES-buffer (200 mM Tris pH8, 20 % sucrose; 50 mM EDTA). The cells are then incubated for 10 min on ice and vortexed regularly. Subsequently, the mixture is centrifuged at 10,000 rpm for 1.5 min and the supernatant is disgarded. The pellet is quickly taken up in 1 ml ice cold distilled water. After 10 min on ice and regular vortexing the mixture is centrifuged at 14,000 rpm for 2 min and the resulting supernatant is recuperated as a periplasmic fraction. This material was compared to pelleted cells and material leaking into the medium, by means of SDS-page and detection with anti-His or anti-myc. This demonstrates that part of the expressed 16 kD SCA protein is indeed present in the periplasm and that part is leaking out into the medium (results not shown). The material present in the medium was dialysed against PBS and analyzed by means of SDS-page and detection with anti-CTLA4 or anti -myc (see figure 7) . The 16 kD SCA-protein is clearly present. The 83 kD band is probably due to a fusion of the SCA protein with the gene 3 coat protein of the phage.

Example 5: The soluble, bacterially expressed, non-glycosylated SCA-molecule binds B7.1 and B7.2 on the membrane

The soluble SCA-molecule is tested for binding capacity to the human B7.1 and B7.2 molecules. The EBV-transformed B cell line (RPMI 8866, 0.5-1x 10^5 cells/sample), which strongly expresses B7.1 and B7.2, is incubated for 20 min at 4°C with the soluble SCA-molecule. After washing twice in RPMI 1640 supplemented with 10% FCS, the cells are incubated for another 20 min at 4°C with mouse anti-cMyc mAb (1µg/cell pellet) conjugated to biotine. After washing twice in RPMI 1640 supplemented with 10% FCS, the cells are incubated for another 20 min at 4°C with Streptavidin conjugated to PE (Phycoerythrine). The cells are then washed twice in RPMI 1640 supplemented with 10% FCS and finally suspended in PBS supplemented with 1% BSA and 0.1% NaN_3 and analyzed with a FACScan flow cytometer (Becton Dickinson). The specific binding of the SCA-molecule is expressed as the mean fluorescent intensity in arbitrary units. The results show that soluble, non-glycosylated SCA binds, in a dose-dependent way, to the RPMI 8866 cells which express B7.1 and B7.2 (results not shown).

Example 6 : Construction of SCA- RGD repertoire for display on filamentous phage

The well characterised RGD containing sequence that binds to integrins is inserted such that it replaces the CDR3 loop of CTLA-4. Sequences are also introduced such that they replace the A1/S1 loop and the S2/A2 loops of CTLA-4 (see example 16 and 17). These loops are situated at the side of the molecule opposite to the CDR sequences which mediate the binding of CTLA-4 to its natural ligands.

A fragment corresponding to CTLA-4 (Figure 9) is digested from pCES1 as an ApaL1 /Not1 fragment. The template CTLA-4 DNA is then subjected to PCR with the primers oligo#1 and oligo#2 (Figure 8). The PCR product is then purified and reamplified with oligo#1 and oligo#3. Fragment DNA is recovered and digested with ApaL1 and Not1 and cloned into pCES1 for phage display (Figure 9). In this way, a library of 1.3 x 10^7 independant clones can be obtained.

Example 7 : Selection of repertoire on recombinant human αvβ3 integrin

Three rounds of selection are performed on recombinant αvβ3 integrin coated onto immunotubes at 10μg/ml (Figure 10 a and b). Polyclonal phage ELISA is then performed to monitor the selection using input phage for 1 2 and 3 rounds of selection. As can be seen initially the repertoire has reactivity to B7-1Ig representing the probable presence of wild-type CTLA-4 in the repertoire. As the selection progresses this disappears and gives way to a positive signal on integrin (Figure 11).

Selected clones are tested for binding to antigen in phage ELISA as in (Marks et al., 1991). Phage ELISA is performed on αvβ3 integrin (1μg/ml), B7-1Ig (1μg/ml) and BSA (1μg/ml). In all cases all clones are negative on B7-1Ig and BSA. (Figure 10).

Examples 8 : Sequences of selected clones positive in phage ELISA

Positive clones are sequenced and some sequences are shown in (Figure 12) by way of example. Also, another pointmutation is shown to be present in the CDR2 loop if the CTLA-4 molecule is taken as a reference, further demonstrating another permissive site.

Example 9 : Binding of CTLA-4 RGD molecules in soluble ELISA

Soluble CTLA-4 RGD fragments are prepared as in Marks et al. (1991) and tested for binding to αvβ3 integrin (1μg/ml) either as monomeric reagent or cross linked via 9E10 to generate avidity. As can be seen, clones are positive in soluble ELISA (Figure 13). 14 out of the 26 clones that are positive in phage ELISA (Figure 10) are also positive insoluble ELISA. A positive control is the mouse monoclonal antibody to αvβ3 integrin, LM609 (Brooks et al., 1994). A negative control is a non binding clone as identified by phage ELISA.

Example 10 : ELISA with antibodies to CTLA-4

To test if the integrity of the immunoglobulin fold of the Sca molecule is maintained after introduction of the RGD sequence into the CDR3, an ELISA can be performed using mouse polyclonal anti CTLA-4 serum and a monoclonal antibody to CTLA-4 (IGH310). This antibody is supposed to compete with B7-1/B7-2 so it has partial epitope overlap with the MYPPPY sequence of wild-type CTLA-4.

5×10^{11} phage are coated directly in PBS overnight at 4°C and are challenged with either mouse polyclonal anti-CTLA-4 or mouse monoclonal anti-CTLA-4. As can be seen, clones 7B, 8H and 11C and wild-type CTLA-4 are all positive with both antibodies. The signal obtained with the monoclonal antibody is greater than that obtained with the polyclonal serum. Furthermore the signal obtained with the RGD presenting clones in CDR3 is greater than that obtained with CTLA-4. This indicates that the general structure of SCA molecule remains intact after introduction of foreign sequences into CDR3 due to the positive signal obtained with the conformation dependant monoclonal anti-CTLA-4. (Figure 14)

Example 11 : FACS analysis of binding to HUVECS cells

Approximately 80,000 HUVECS cells/sample are cultured and are tested in FACS for binding. As can be seen a positive signal is seen with the positive clone . Whereas with both empty pCES1 vector and wild-type CTLA-4 phage no shift can be observed. (Figure 15). This example demonstrates that the selected RGD-containing scaffolds not only bind to integrin but also exhibit substantial specificity.

Example 12 : Construction of SCA-E

The gene coding for SCA, preceded by the pelB signal sequence and followed by a combined His6/c-myc tag, is available in the phage display vector pCES1CTLA4 (see example 3). This vector is cut with NotI and ScaI and a 2183 bp fragment is isolated and ligated with a 2654 bp NotI/ScaI fragment originating from vector pCANTAB5E (Amersham Pharmacia, Uppsala, Sweden). The resulting vector pCES1CTLA4E differs

from the previously described vector pCES1CTLA4 in that the CTLA4 molecule now carries a C-terminal E-tag in stead of a double c-myc/His6 tag. The correct sequence of the vector is then verified by DNA sequencing.

Example 13 : Display of the recombinant SCA-E protein on the surface of filamentous phage

Phage displaying SCA-E protein on their surface are prepared as described in example 2 (Marks et al., 1991). To demonstrate binding of the SCA-E protein to B7.1Ig, B7.2Ig and an anti E-tag antibody (Amersham Pharmacia, Uppsala, Sweden), phages are titrated and a serial dilution is made to test for binding in ELISA to these molecules. Strong signals are obtained showing binding to both B7.1Ig and B7.2Ig and the anti E-tag antibody. This demonstrates that SCA-E has the proper conformation and remains functional as a monomer.

Example 14 : Insertion of an integrin binding sequence (RGD) in the loop formed by β-sheets A1 and S1 in SCA-E

Using plasmid pCES1CTLA4E (see example 3) as a template, the DNA coding for the integrin binding sequence (RGD) is inserted in the loop formed by β-sheets A1 and S1 using overlap extension PCR (Ho et al., 1989). In the A1/S1 loop, consisting of residues SSRGI (see figure 2), the RGD sequence is inserted between the SS and the RGI residues. On both sides, the RGD sequence is flanked by a G4S linker sequence. This insertion can be accomplished by performing two PCR reactions using pCES1CTLA4E plasmid DNA as a template. A first PCR, using primers Nos. 11209 and 11071, amplifies the 5' end of the CTLA4E gene including the RGD sequence. The second PCR, using primers Nos. 11072 and 11210, amplifies the 3' end of the CTLA4E gene including the RGD sequence. Both PCR fragments are subsequently gelpurified and an additional PCR is performed on the mixture of both PCR fragments using primers Nos. 11209 and 11210. The resulting PCR fragment is gelpurified, cut with ApaLI and NotI, and inserted into the ApaLI/NotI opened vector pCES1CTLA4E. The resulting vector pCES1CTLA4E-A1/S1RGD is verified by DNA sequencing.

Primer sequences:

No. 11071 (SEQ ID 236):
5'-GTCACCACGAGAGCCACCGCCACCGCTGCTGGCCAGTACCACAG-3'
No. 11072 (SEQ ID 237):
5'-CGTGGTGACGGTGGCGGTGGCTCTCGAGGCATCGCCAGCTTTG-3'
No. 11209 (SEQ ID 238):
5'-GATTACGCCAAGCTTTGGAGC-3'
No. 11210 (SEQ ID 239):
5'-ATGCGGCCCCATTCAGATC-3'

Example 15 : Insertion of a His6 tag (HHHHHH) in the loop formed by β-sheets A1 and S1 in SCA-E

Using plasmid pCES1CTLA4E (see example 3) as a template, the DNA coding for a His6 tag (HHHHHH) is inserted in the loop formed by β-sheets A1 and S1 using overlap extension PCR (Ho et al., 1989). In the A1/S1 loop, consisting of residues SSRGI (see figure 2), the HHHHHH sequence is inserted between the SS and the RGI residues. On both sides, the HHHHHH sequence is flanked by a G4S linker sequence. This insertion can be accomplished by performing two PCR reactions using pCES1CTLA4E plasmid DNA as a template. A first PCR, using primers Nos. 11209 and 11069, amplifies the 5' end of the CTLA4E gene including the HHHHHH sequence. The second PCR, using primers Nos. 11070 and 11210, amplifies the 3' end of the CTLA4E gene including the HHHHHH sequence. Both PCR fragments are subsequently gelpurified and an additional PCR is performed on the mixture of both PCR fragments using primers Nos. 11209 and 11210. The resulting PCR fragment is gelpurified, cut with ApaLI and NotI, and inserted into the ApaLI/NotI opened vector pCES1CTLA4E. The resulting vector pCES1CTLA4E-A1/S1His6 is verified by DNA sequencing.

Primer sequences:

No. 11069 (SEQ ID 240):

5'-GTGGTGATGGTGATGGTGAGAGCCACCGCCACCGCTGCTGGCCAGTACC ACAG-3'

No. 11070 (SEQ ID 241):

5'-CACCATCACCATCACCACGGTGGCGGTGGCTCTCGAGGCATCGCCAGCT TTG-3'

No. 11209 (SEQ ID 242):

5'-GATTACGCCAAGCTTTGGAGC-3'

No. 11210 (SEQ ID 243):

5'-ATGCGGCCCCATTCAGATC-3'

Example 16 : Insertion of an integrin binding sequence (RGD) in the loop formed by β-sheets S2 and A2 in SCA-E

Using plasmid pCES1CTLA4E (see example 3) as a template, the DNA coding for the integrin binding sequence (RGD) is inserted in the loop formed by β-sheets S2 and A21 using overlap extension PCR (Ho et al., 1989). In the S2/A2 loop, consisting of residues DS (see figure 2), the RGD sequence is inserted between the D and the S residue. No flanking sequences are present. This insertion can be accomplished by performing two PCR reactions using pCES1CTLA4E plasmid DNA as a template. A first PCR, using primers Nos. 11209 and 11079, amplifies the 5' end of the CTLA4E gene including the RGD sequence. The second PCR, using primers Nos. 11080 and 11210, amplifies the 3' end of the CTLA4E gene including the RGD sequence. Both PCR fragments are subsequently gelpurified and an additional PCR is performed on the mixture of both PCR fragments using primers Nos. 11209 and 11210. The resulting PCR fragment is gelpurified, cut with ApaLI and NotI, and inserted into the ApaLI/NotI opened vector pCES1CTLA4E. The resulting vector pCES1CTLA4E-S2/A2RGD is verified by DNA sequencing.

Primer sequences:

No. 11079 (SEQ ID 244):
5'-CACCTGGCTGTCACCACGGTCAGCCTGCCGAAGCACTG-3'
No. 11080 (SEQ ID 245):
5'-CAGGCTGACCGTGGTGACAGCCAGGTGACTGAAGTCTGTGC-3'
No. 11209 (SEQ ID 246):
5'-GATTACGCCAAGCTTTGGAGC-3'
No. 11210 (SEQ ID 247):
5'-ATGCGGCCCCATTCAGATC-3'

Example 17 : Insertion of a His6 tag (HHHHHH) in the loop formed by β-sheets S2 and A1 in SCA-E

Using plasmid pCES1CTLA4E (see example 3) as a template, the DNA coding for a His6 tag (HHHHHH) is inserted in the loop formed by β-sheets S2 and A2 using overlap extension PCR (Ho et al., 1989). In the S2/A2 loop, consisting of residues DS (see figure 2), the HHHHHH sequence is inserted between the D and the S residue. No flanking sequences are present. This insertion can be accomplished by performing two PCR reactions using pCES1CTLA4E plasmid DNA as a template. A first PCR, using primers Nos. 11209 and 11075, amplifies the 5' end of the CTLA4E gene including the HHHHHH sequence. The second PCR, using primers Nos. 11077 and 11210, amplifies the 3' end of the CTLA4E gene including the HHHHHH sequence. Both PCR fragments are subsequently gelpurified and an additional PCR is performed on the mixture of both PCR fragments using primers Nos. 11209 and 11210. The resulting PCR fragment is gelpurified, cut with ApaLI and NotI, and inserted into the ApaLI/NotI opened vector pCES1CTLA4E. The resulting vector pCES1CTLA4E-S2/A2His6 is verified by DNA sequencing.

Primer sequences:

No. 11075 (SEQ ID 248):

5'-GTGGTGATGGTGATGGTGGTCAGCCTGCCGAAGCACTG-3'

No. 11077 (SEQ ID 249):

5'-CACCATCACCATCACCACAGCCAGGTGACTGAAGTCTGTGC-3'

No. 11209 (SEQ ID 250):

5'-GATTACGCCAAGCTTTGGAGC-3'

No. 11210 (SEQ ID 251):

5'-ATGCGGCCCCATTCAGATC-3'

Example 18 : SCA-E A1/S1 His6 and SCA-E S2/A2 His6 displayed on phage are able to bind coated anti His6-tag antibodies in ELISA

Phage displaying SCA-E A1/S1 His6 or SCA-E S2/A2 His6 protein on their surface were prepared as described by Marks et al. (1991). To demonstrate binding of the SCA-E A1/S1 His6 or SCA-E S2/A2 His6 protein to an anti His6-tag antibody (BabCO, Richmond, CA, USA), phages are titrated and a serial dilution is made to test for binding in ELISA to these molecules. Significant signals were obtained showing binding to the anti His6-tag antibody.

Example 19 : SCA-E A1/S1 His6 and SCA-E S2/A2 His6 displayed on phage are able to bind Ni^{2+} coated plates in ELISA

Phage displaying SCA-E A1/S1 His6 or SCA-E S2/A2 His6 protein on their surface were prepared as described by Marks et al. (1991). To demonstrate binding of the SCA-E A1/S1 His6 or SCA-E S2/A2 His6 protein to Ni-NTA HisSorbTM Plates (Qiagen GmbH, Hilden, Germany), phages are titrated and a serial dilution is made to test for binding to these plates. Significant signals were obtained showing binding to the Ni-NTA coated plates.

Example 20 : SCA-E A1/S1 His6 and SCA-E S2/A2 His6 displayed on phage are still able to bind coated recombinant B7.1Ig and B7.2Ig in ELISA

Phage displaying SCA-E protein on their surface are prepared as described by

Marks et al. (1991). To demonstrate binding of the SCA-E A1/S1 His6 or SCA-E S2/A2 His6 proteins to B7.1Ig and B7.2Ig, phages are titrated and a serial dilution is made to test for binding in ELISA to these molecules. Strong signals can be obtained demonstrating binding to both B7.1Ig and B7.2Ig. This demonstrates that the SCA-E E A1/S1 His6 or SCA-E S2/A2 His6 proteins have the proper conformation and remain functional as a monomer. Furthermore, it demonstrates that both the A1/S1 and S2/A2 loops are permissive sites for insertion of peptide sequences without destroying the overall scaffold conformation.

Example 21 : SCA-E A1/S1 RGD and SCA-E S2/A2 RGD displayed on phage are still able to bind coated recombinant B7.1Ig and B7.2Ig in ELISA

To test for the integrity of CTLA-4 derivatives in which foreign sequences are introduced into the loops A1/S1 and S2/A2, a phage ELISA is performed on antigens αvβ3 integrin (1μg/ml), B7-1Ig and B7-2Ig (1μg/ml) and BSA (1μg/ml). As can be seen AB' KL10 shows a positive signal on B7-1Ig and B7-2Ig suggesting that despite the introduction of the sequence SGGGS into the A1/S1 loop, the functional integrity of the molecule is maintained (Figure 14), proving the concept of the invention that the molecule of the presdent invention can function as a scaffold. The integrity of these molecules is further verified as a weak binding in phage ELISA as can be seen on anti CTLA-4 (Figure 16).

LIST OF REFERENCES

-**Abedi, M. R., Caponigro, G., and Kamb, A.** (1998) Green fluorescent protein as a scaffold for intracellular presentation of peptides, Nucleic Acids Res. 26: 623-30

-**Amit,A.G., Mariuzza,R.A.,Phillips,S.E.V., and Poljak,R.J.** (1986) Three dimensional structure of an antigen-antibody complex at 2.8Å resolution. Science 233: 747

-**Amzel,L.M. and Poljak,R.J.** (1979) Three dimensional structure of immunoglobulins. Ann.Rev.Biochem. 48: 961

-**Atherton & Shepard** (1989) Solid phase peptide synthesis. IRL Press, Oxford.

-**Barbas, C. F. d., Languino, L. R., and Smith, J. W.** (1993) High-affinity self-reactive human antibodies by design and selection: targeting the integrin ligand binding site, Proc Natl Acad Sci U S A. 90: 10003-7

-**Bianchi,E.,Sollazzo,M.,Tramontano,A.,Pessi,A.** (1993) Affinity purification of a difficult sequence protein. Int.J.Peptide Protein Res. 42: 93

-**Bianchi,E,Venturini,S.,Pessi,A.,Tramontano,A.,Sollazzo,M.** (1994) High level expression and rational mutagenesis of a designed protein, the minibody: from an insoluble to a soluble protein.

-**Brooks, P. C., Montgomery, A. M., Rosenfeld, M., Reisfeld, R. A., Hu, T., Klier, G., and Cheresh, D. A.** (1994) Integrin alpha v beta 3 antagonists promote tumor regression by inducing apoptosis of angiogenic blood vessels, Cell. 79: 1157-64

-**Cannon E.L., Ladner R.C., McCoy D.** (1996) Phage-display technology. IVD Technology Nov./Dec., Canon Communications.

-**Choo, Y. and Klug, A.** (1995) Designing DNA-binding proteins on the surface of filamentous phage, Curr Opin Biotechnol. 6: 431-6

-**Colas, P., Cohen, B., Jessen, T., Grishina, I., McCoy, J., and Brent, R.** (1996) Genetic selection of peptide aptamers that recognize and inhibit cyclin- dependent kinase 2, Nature. 380: 548-50

-**Hamers-Casterman C., Atarhouch T., Muyldermans S., Robinson G., Hamers C., Songa E.B., Bendahman N., Hamers R.** (1993) Naturally occuring antibodies devoid of light chains. Nature 363: 446-448

-**Ho. S., Hunt H., Horton H., Pullen J. & Pease L.** (1989) Site-directed mutagenesis by

overlap extension using the polymerase chain reaction. Gene, 77, pp. 51-59.

- **Houbenweyl** (1974) Methode der organischen chemie, vol. 15, I & II (ed. Wunch E). Thieme, Stuttgart.

-**Houston, M. E., Jr., Wallace, A., Bianchi, E., Pessi, A., and Hodges, R. S.** (1996) Use of a conformationally restricted secondary structural element to display peptide libraries: a two-stranded alpha-helical coiled-coil stabilized by lactam bridges, J Mol Biol. 262: 270-82,

-**Knapp, A., Degenhardt, T., and Dodt, J. Hirudisins.** (1992) Hirudin-derived thrombin inhibitors with disintegrin activity, J Biol Chem. 267: 24230-4

-**Koide, A., Bailey, C. W., Huang, X., and Koide, S.** (1998) The fibronectin type III domain as a scaffold for novel binding proteins, J Mol Biol. 284: 1141-51

-**Ku, J. and Schultz, P. G.** (1995) Alternate protein frameworks for molecular recognition, Proc Natl Acad Sci U S A. 92: 6552-6

-**Kubinyi H.** (1995). Strategies and recent technologies in drug discovery. Pharmazia 50:647-662.

-**Ladner R.** (1995) Constrained peptides as binding entities.TibTech 13:426-430.

-**Maclennan J.** (1995) Engineering microprotein ligands for large-scale affinity purification. Biotechnology 13: 1181-1183.

-**Maniatis T., Fritsch E., Sambrook J.** (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

-**Marks JD, Hoogenboom HR, Bonnert TP, McCafferty J, Griffiths AD, Winter G** (1991) By-passing immunisation. Human antibodies from V-gene libraries displayed on phage, J. Mol. Biol.; 222(3), pp. 581-97

-**Martin, F., Toniatti, C., Salvati, A. L., Venturini, S., Ciliberto, G., Cortese, R., and Sollazzo, M.** (1994) The affinity-selection of a minibody polypeptide inhibitor of human interleukin-6, Embo J. 13: 5303-9

-**McConnell, S. J. and Hoess, R. H.** (1995) Tendamistat as a scaffold for conformationally constrained phage peptide libraries, J Mol Biol. 250: 460-70

-**Metzler W.J., Bajorath J., Fenderson W., Shaw S., Constantine K., Naemura J., Leytze G., Peach R., Lavoie T., Mueller L., Linsley P.** (1997) Solution structure of

human CTLA-4 and delineation of a CD80/CD86 binding site conserved in CD28. Nat. Struct. Biol. 4: 527-531.

-**Neu,H.C. and Heppel,L.A.** (1965) J.Biol.Chem. 240: 3685

-**Nord, K., Gunneriusson, E., Ringdahl, J., Stahl, S., Uhlen, M., and Nygren, P.** (1997) A. Binding proteins selected from combinatorial libraries of an alpha- helical bacterial receptor domain, Nat Biotechnol. 15: 772-7

-**Pasqualini, R., Koivunen, E., and Ruoslahti, E.** (1997) Alpha v integrins as receptors for tumor targeting by circulating ligands [see comments], Nat Biotechnol. 15: 542-6

-**Pasqualini, R., Koivunen, E., and Ruoslahti, E.** (1995) A peptide isolated from phage display libraries is a structural and functional mimic of an RGD-binding site on integrins, J Cell Biol. 130: 1189-96

-**Pessi, A., Bianchi, E., Crameri, A., Venturini, S., Tramontano, A., and Sollazzo, M.** (1993) A designed metal-binding protein with a novel fold [see comments], Nature. 362: 367-9

-**Raag,R., and Whitlow,M.** (1995) Single-chain Fvs. FASEB J. 9:73

-**Reiter,Y.,Brinkmann,U.,Webber,K.,Jung,S.-H.,Lee,B.K. and Pastan,I** (1994) Engineering interchain disulfide bonds into conserved framework regions of Fv fragments: improved biochemical characteristics of recombinant immunotoxins containing disulfide- stabilized Fv. Protein Eng. 7: 697

-**Rottgen, P. and Collins, J.** (1995) A human pancreatic secretory trypsin inhibitor presenting a hypervariable highly constrained epitope via monovalent phagemid display, Gene. 164: 243-50

-**Smith, J. W., Hu, D., Satterthwait, A., Pinz-Sweeney, S., and Barbas, C. F., 3rd** (1994) Building synthetic antibodies as adhesive ligands for integrins, J Biol Chem. 269: 32788-95

-**Smith, G. P., Patel, S. U., Windass, J. D., Thornton, J. M., Winter, G., and Griffiths, A. D.** (1998) Small binding proteins selected from a combinatorial repertoire of knottins displayed on phage, J Mol Biol. 277: 317-32

-**Williams A.F.** (1987) A year in the life of the immunoglobulin superfamily. Immunol.Today 8: 298

CLAIMS

1. A scaffold composed of a single-chain polypeptide having the following structural properties:

- it contains at least two cysteine residues which form at least one disulphide bond, and
- it possesses less than 10% of alpha helical conformation, and
- it contains at least six β-strands (S1, S2, S3, S4, S5, S6) which:
  - are connected by amino acid loops of variable conformation and lenght according to the following topology: S1-S2-S3-S4-S5-S6, and wherein at least one of said loops is different for at least three, preferably four and more preferably five contiguous amino acid residues by way of substitution, deletion or insertion when compared to the naturally occurring CTLA-4 molecule,
- it forms two β-sheets, one formed by S1/S4/S3 and one formed by S6/S5/S2 wherein the symbol "/" denotes the hydrogen bonding interactions between two spatially adjacent β strands, which are each characterized by an anti-parallel arrangement of said β strands and which are packed onto eachother so that they form a β sandwich architecture.

2. A scaffold according to claim 1 which:

- contains maximum three additional β strands (A1, A2 and A3), and
- has the following said topology: A1-S1-S2-A2-A3-S3-S4-S5-S6 wherein the β strands are connected by amino acid loops of variable conformation and lenght, and wherein at least one of said loops is different for at least three contiguous amino acid residues by way of substitution, deletion or insertion when compared to the naturally occurring CTLA-4 molecule,
- possesses the following said β-sheets: A1/S1/S4/S3 and S6/S5/S2/A2/A3.

3. A scaffold according to claims 1 and 2 wherein:

-A1 is the amino acid sequence AQPAVVLA (SEQ ID 1) or any functionally equivalent derivative of said sequence,

-S1 is the amino acid sequence ASFPVEY (SEQ ID 2) or any functionally equivalent derivative of said sequence,

-S2 is the amino acid sequence EVRVTVLRQA (SEQ ID 3) or any functionally equivalent derivative of said sequence,

-A2 is the amino acid sequence QVTEVCAA (SEQ ID 4) or any functionally equivalent derivative of said sequence,

-A3 is the amino acid sequence TYMMGNELTFLDDS (SED ID 5) or any functionally equivalent derivative of said sequence,

-S3 is the amino acid sequence ICTGTSS (SEQ ID 6) or any functionally equivalent derivative of said sequence,

- S4 is the amino acid sequence QVNLTIQ (SEQ ID 7) or any functionally equivalent derivative of said sequence,

- S5 is the amino acid sequence GLYICKVE (SEQ ID 8) or any functionally equivalent derivative of said sequence,

- S6 is the amino acid sequence GIGNGTQIY (SEQ ID 9) or any functionally equivalent derivative of said sequence.

4.A scaffold according to claims 1 to 3 which does not bind to B7.1 or B7.2.

5. A scaffold according to claims 1 to 4 wherein at least two, preferably three and more preferably four of said loops is different for at least three, preferably at least four and more preferably at least five contiguous amino acid residues by way of substitution, deletion or insertion when compared to the naturally occurring CTLA-4 molecule.

6. A scaffold according to claims 1 to 5 wherein said amino acid loops comprise a binding domain for a receptor or antigen.

7. A scaffold according to claims 1 to 5 wherein said amino acid loops comprise two

separate binding domains for receptors and/or antigens.

8. A scaffold according to claim 7 wherein the two separate binding domains are binding to identical molecules, preferably disease specific antibodies, thereby providing a tool for agglutination of such antibodies.

9. A scaffold according to claim 7 wherein one of the binding domains is capable of binding to red blood cell surface proteins, preferably glycophorin, thereby providing a tool for blood agglutination tests.

10. A scaffold according to claims 1 to 9 wherein at least one of said amino acid loops provides for a spacer or anchor to allow for immobilisation on a support or carrier.

11. A scaffold according to claims 1 to 10 wherein at least one of said amino acid loops provides a His-tag.

12. A scaffold according to claims 1 to 11 wherein at least one of said amino acid loops provides an RGD-sequence that has integrin binding capacity.

13. A scaffold according to claims 1 to 12 wherein at least one of said amino acid loops provides for a CD3-like binding domain that can activate cytotoxic T lymphocytes.

14. A scaffold according to claims 1 to 13 wherein at least one of said binding sites comprises loops that can act as a chelator for rare earth metals.

15. A scaffold according to claims 1 to 14 wherein said single-chain polypeptide is unglycosylated.

16. A scaffold according to claims 1 to 14 wherein said single-chain polypeptide is glycosylated.

17. Use of a single-chain polypeptide according to claim 1 as a scaffold.

18. Use of a single-chain polypeptide according to claim 2 as a scaffold.

19. Use of a single-chain polypeptide according to claims 1 or 2 to restrain peptides with randomized amino acid sequences between said β-strands within loops.

20 . Use of a single-chain polypeptide according to claims 1 or 2 for the preparation of randomized peptide libraries.

21 . Use of a scafffold according to any of claims 1 to 16 as a diagnostic tool.

22 . Use of a scaffold according to any of claims 1 to 16 as a medicament.

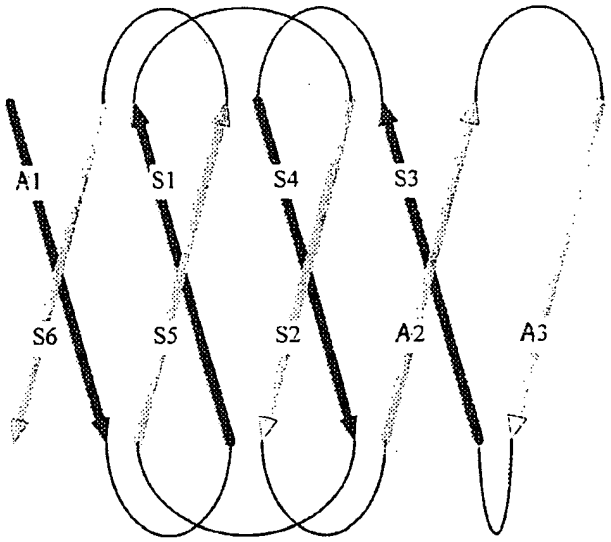


Figure 1

A1

GCA ATG CAC GTG GCC CAG CCT GCT GTG GTA CTG GCC
A M H V A Q P A V V L A
- - - - - -CTLA-4 (Front) - - - - >

S1

AGC AGC CGA GGC ATC GCC AGC TTT GTG TGT
S S R G I A S F V C
Cys21
(Ig)

GAG TAT GCA TCT CCA GGC AAA GCC ACT GAG
E Y A S P G K A T E
+/+ +/+ E31S
Tyr23 Ser25 -/-

S2

GTC CGG GTG ACA GTG CTT CGG CAG GCT GAC
V R V T V L R Q A D>
R33S T35S L37S
-/+ +/+ +/+

A2

AGC CAG GTG ACT GAA GTC TGT GCG GCA ACC
S Q V T E V C A A T
E46S Cys48 T51S
+/- (Non-Ig) +/+

A3

TAC ATG ACG GGG AAT GAG TTG ACC TTC CTA
Y M T G N E L T F L>

S3

GAT GAT TCC ATC TGC ACG GGC ACC TCC AGT
D D S I C T G T S S
Cys68
(Non-Ig)

S4

GGA AAT CAA GTG AAC CTC ACT ATC CAA GGA
G N Q V N L T I Q G>
Asn78
(Glycosylation site)

CTG AGG GCC ATG GAC ACG GGA CTC TAC ATC
L R A M D T G L Y I

S5

TGC AAG GTG GAG CTC ATG TAC CCA CCG CCA
C K V E L M Y P P P>
Cys94 K95S E97 Leu98
(Ig) -/-

Figure 2

S6

TAC TAC CTG GGC ATA GGC AAC GGA ACC CAG
Y Y L G I G N G T Q
Leu106 Asn111
(Glycosylation site)

ATT TAT GTA ATT GAT CCA GAA CCG TGC
I Y V I D P E P C>
<- - - - - - -CTLA-4 (Back) - - - - - - - - - -

Figure 2 - cont'd

CTLA-4 (Front)

*Apa*L1

5' ggcatgcaatgca **GT GCA C** AG GCA ATG CAC GTG GCC CAG CCT GCT 3'

Ala Gln Ala Met His Val Ala Gln Pro Ala

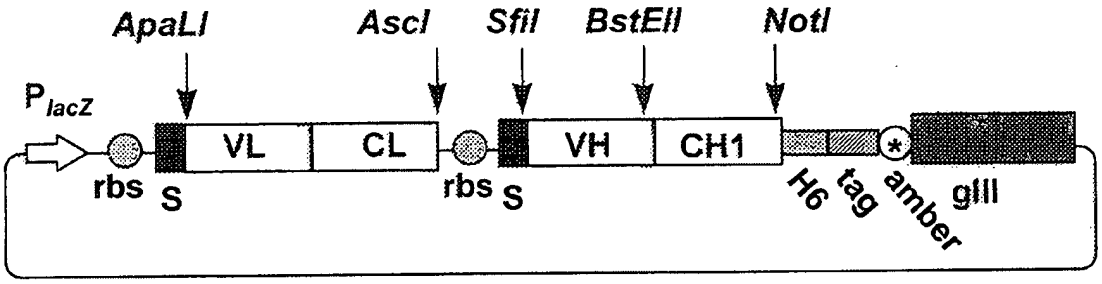
CTLA-4 (Back)

*Not*1

5' gccgcagttgcaa **GC GGC CGC** CGG TTC TGG ATC AAT TAC ATA AAT 3'

Ala Ala Pro Glu Pro Asp Ile Val Tyr Ile

Figure 3



pCES1

Figure 4

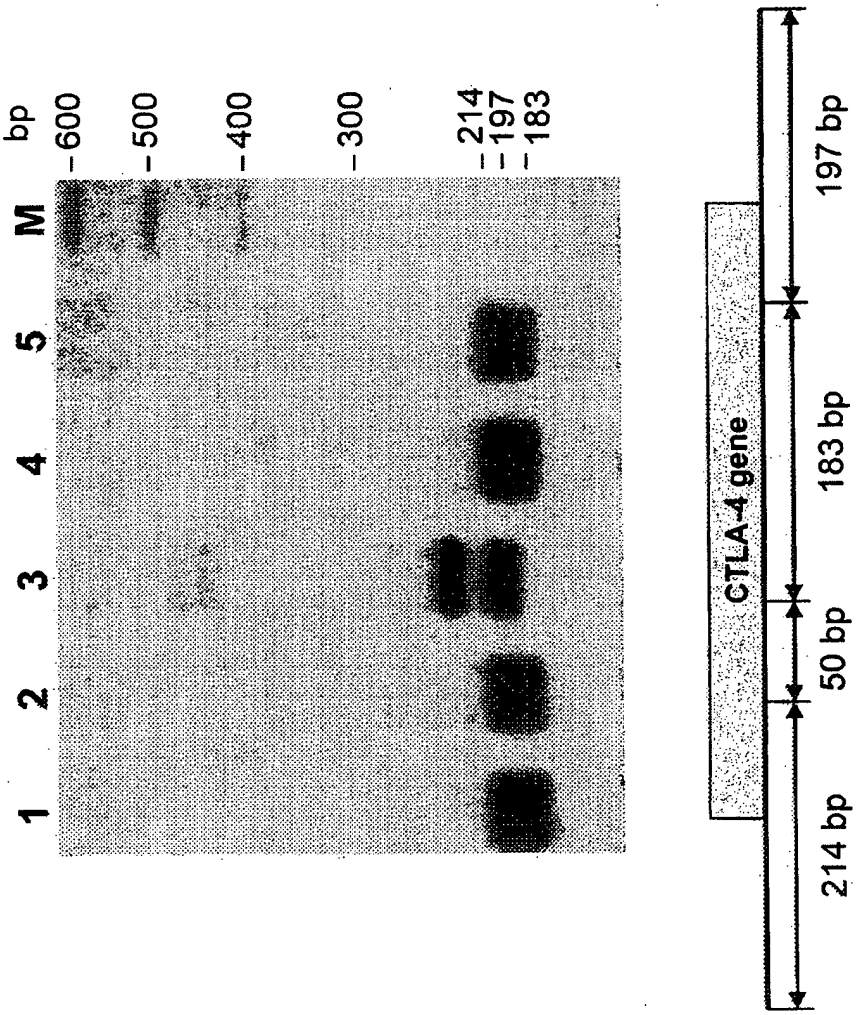

bp
600
500
400
300
214
197
183
M
5
4
3
2
1
CTLA-4 gene
197 bp
183 bp
50 bp
214 bp

Figure 5

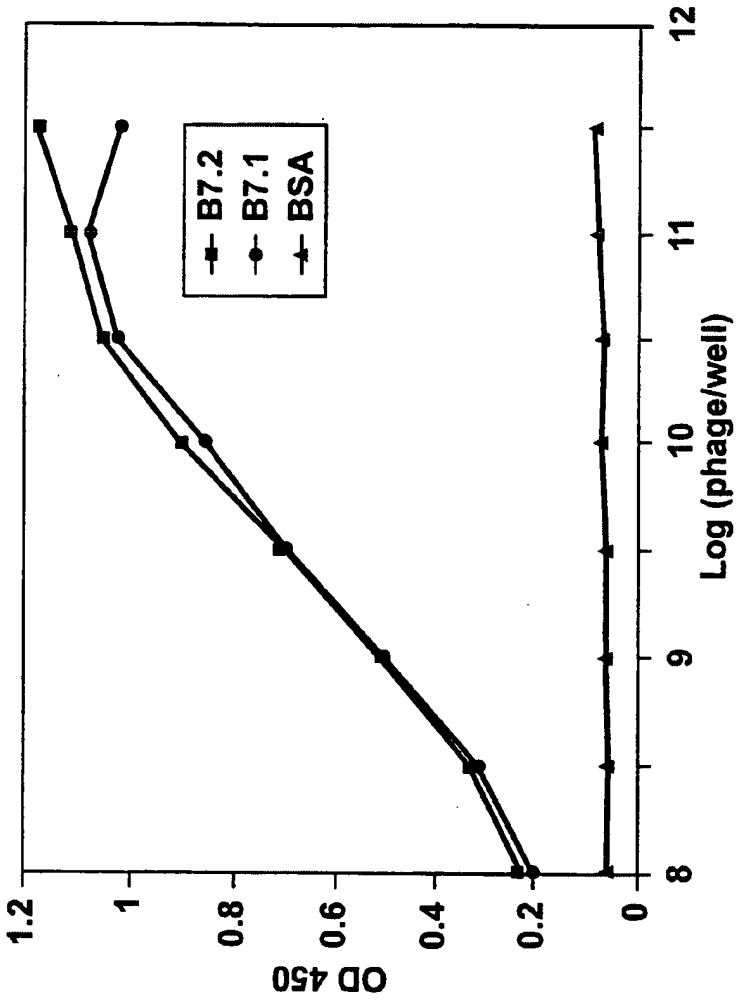


Figure 6

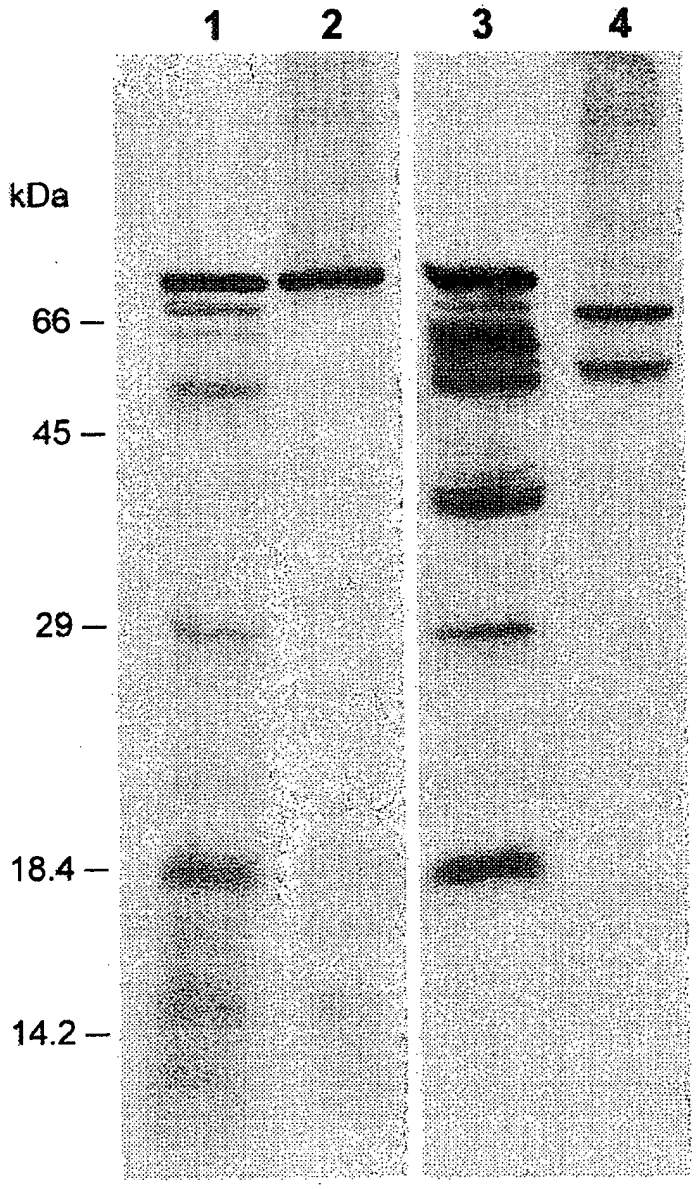


Figure 7

Figure 8

Oligo1

*Apa*L1

5' ggcatgcaatgca**GT GCA C**AG GCA ATG CAC GTG GCC CAG CCT GCT 3'

Ala Gln Ala Met His Val Ala Gln Pro Ala

Oligo 2

```
CTC TAC ATC TGC AAG GTG GAG NNK NNK NNK CGT GGC GAT NNK NNK NNK GGC ATA GGC AAC
GAG ATG TAG ACG TTC CAC CTC NNM NNM NNM GCA CCG CTA NNM NNM NNM CCG TAT CCG TTG
 L   Y   I   C   K   V   E   X   X   X   R   G   D   X   X   X   G   I   G   N>

GGA ACC CAG ATT  3'
CCT TGG GTC TAA  5'
 G   T   Q   I
```

Reamplification oligo 3

```
                                        Not1
GGC AAC GGA ACC CAG ATT TAT GTA ATT GAT CCA GAA CCG GCG GCC GC   3'
CCG TTG CCT TGG GTC TAA ATA CAT TAA CTA GGT CTT GGC CGC CGG CG   5'
 G   N   G   T   Q   I   Y   V   I   D   P   E   P   A   A   A>
```

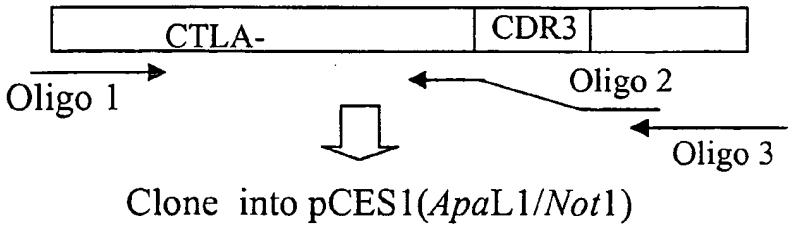


Figure 9

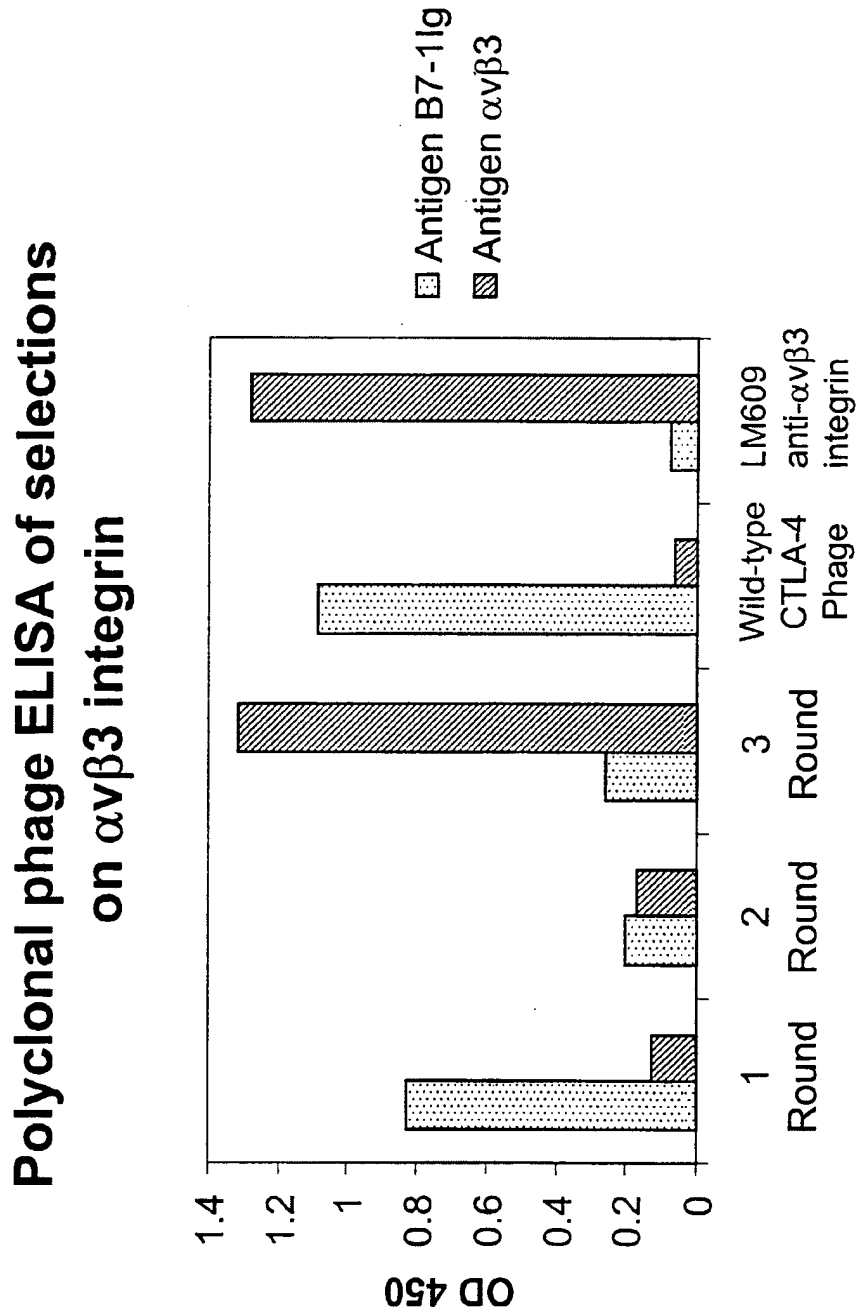

Polyclonal phage ELISA of selections on αvβ3 integrin
Antigen B7-1Ig
Antigen αvβ3
OD 450
1.4
1.2
1
0.8
0.6
0.4
0.2
0
1 Round
2 Round
3 Round
Wild-type CTLA-4 Phage
LM609 anti-αvβ3 integrin

Figure 10a

Round	Input	Output	Ratio	Enrich factor	Number binders
0	1e13	6.9e4	6.9e-9		
1	1.9e13	3.1e5	1.6e-8	2.3x	0
2	1.8e13	2.2e6	1.6e-7	10x	0
3	2.8e12	2e6	7.1e-7	45x	26/40

Figure 10b

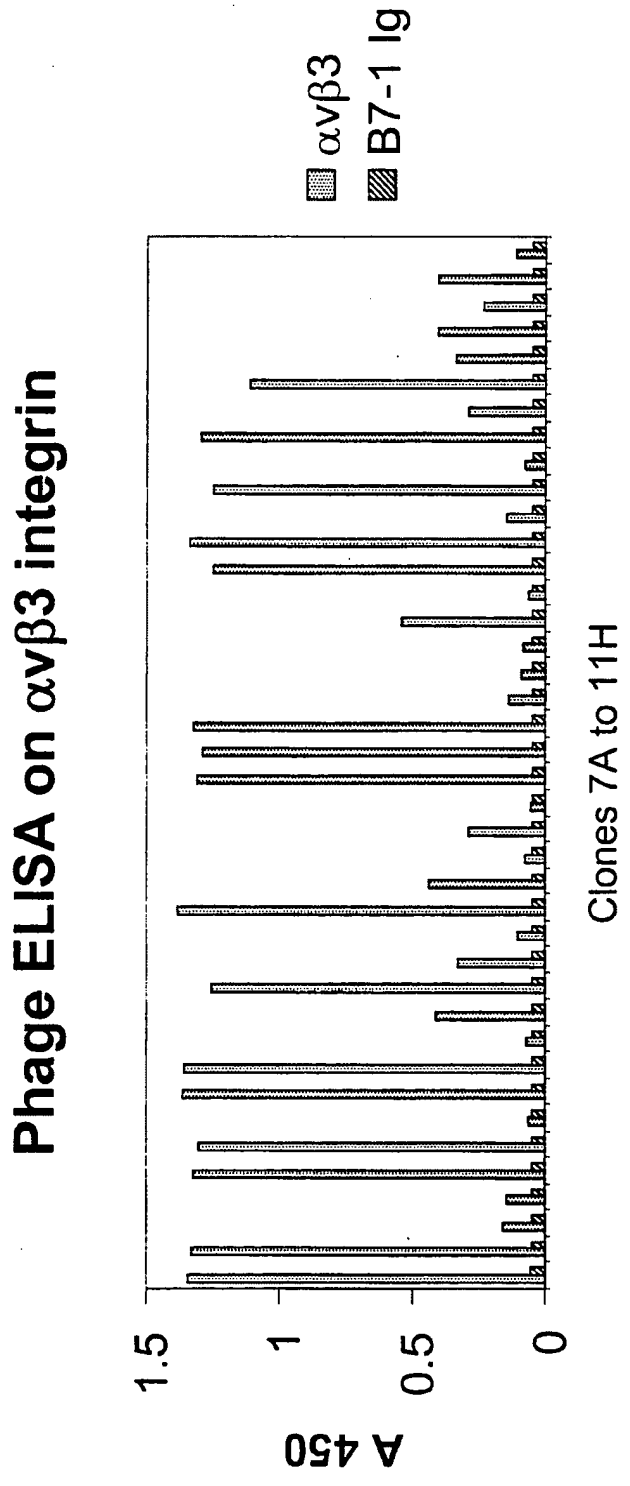

Phage ELISA on αvβ3 integrin
αvβ3
B7-1 Ig
A 450
1.5
1
0.5
0
Clones 7A to 11H

Figure 11

Wt CTLA-4 : ----kveL M Y P P P Y Y L g i g-----
Clone 7B: ----kveF L P RGD D Y P R R g i g--
Clone 8H: ----kveI L D RGD S Y Y g i g---
Clone 11C: ----kveV * G RGD S H P A I g i g---
* is E by nonsense suppression

Figure 12

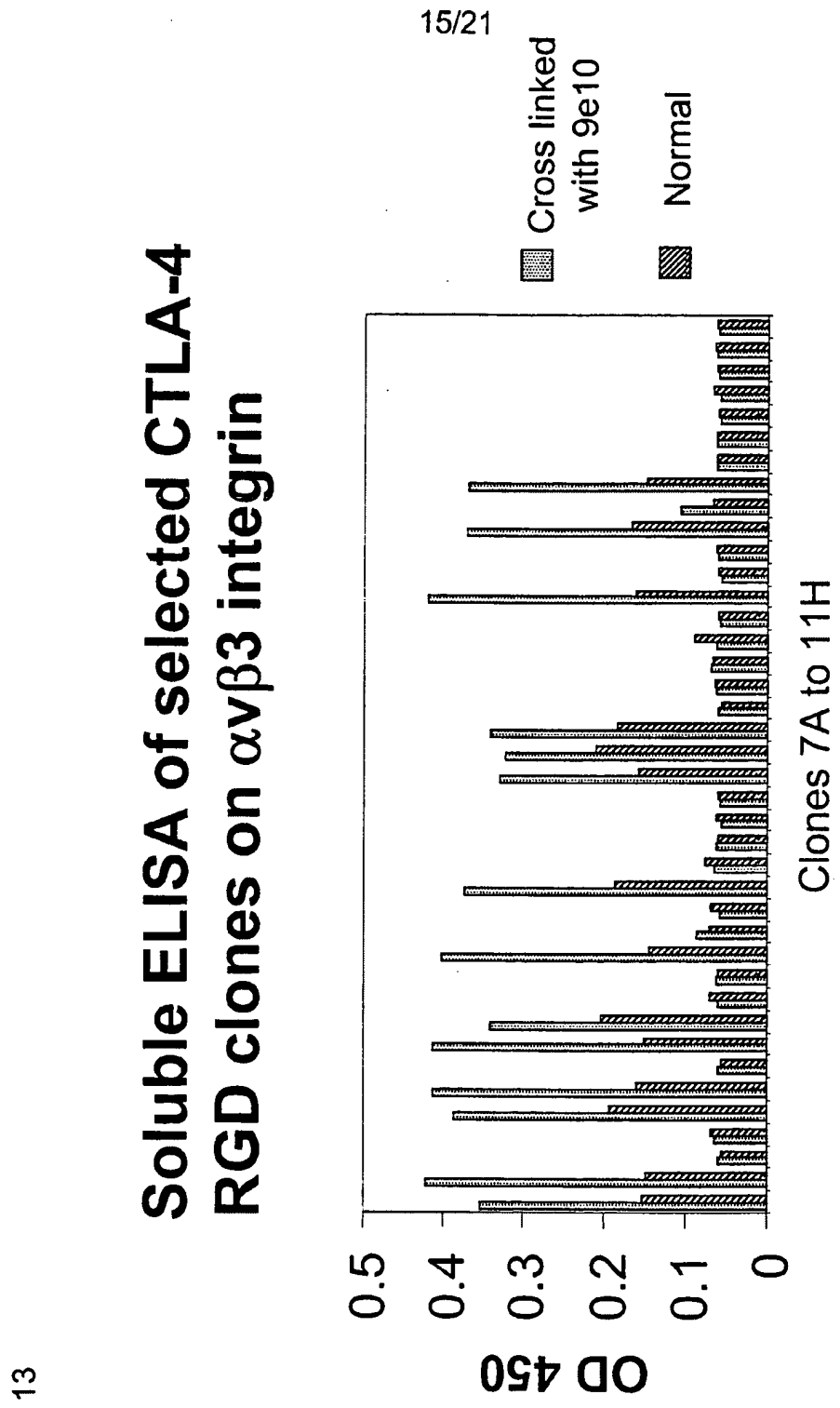

Soluble ELISA of selected CTLA-4
RGD clones on αvβ3 integrin
Cross linked with 9e10
Normal
OD 450
0.5
0.4
0.3
0.2
0.1
0
Clones 7A to 11H

Figure 13

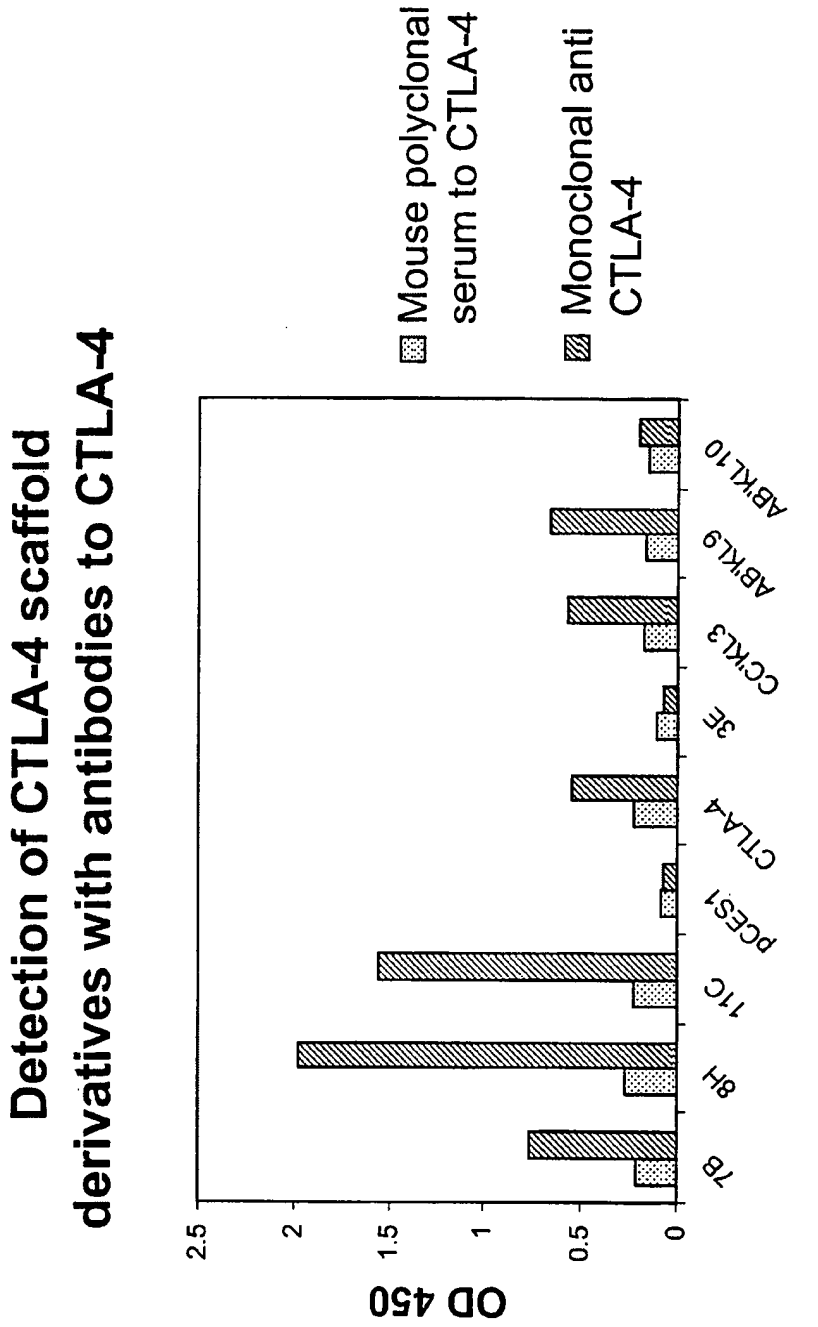

Detection of CTLA-4 scaffold derivatives with antibodies to CTLA-4
Mouse polyclonal serum to CTLA-4
Monoclonal anti CTLA-4
OD 450
2.5
2
1.5
1
0.5
0
7B
8H
11C
PCES1
CTLA4
3E
CCIKL3
ABIKL9
ABIKL10

Figure 14

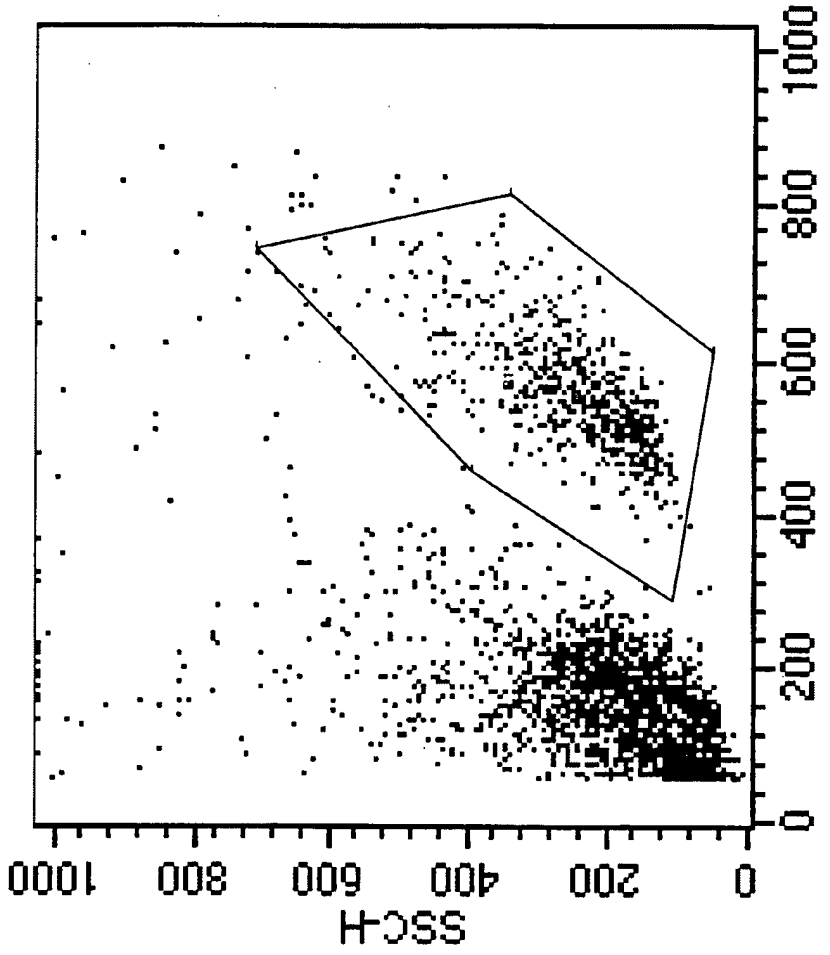


Figure 15A

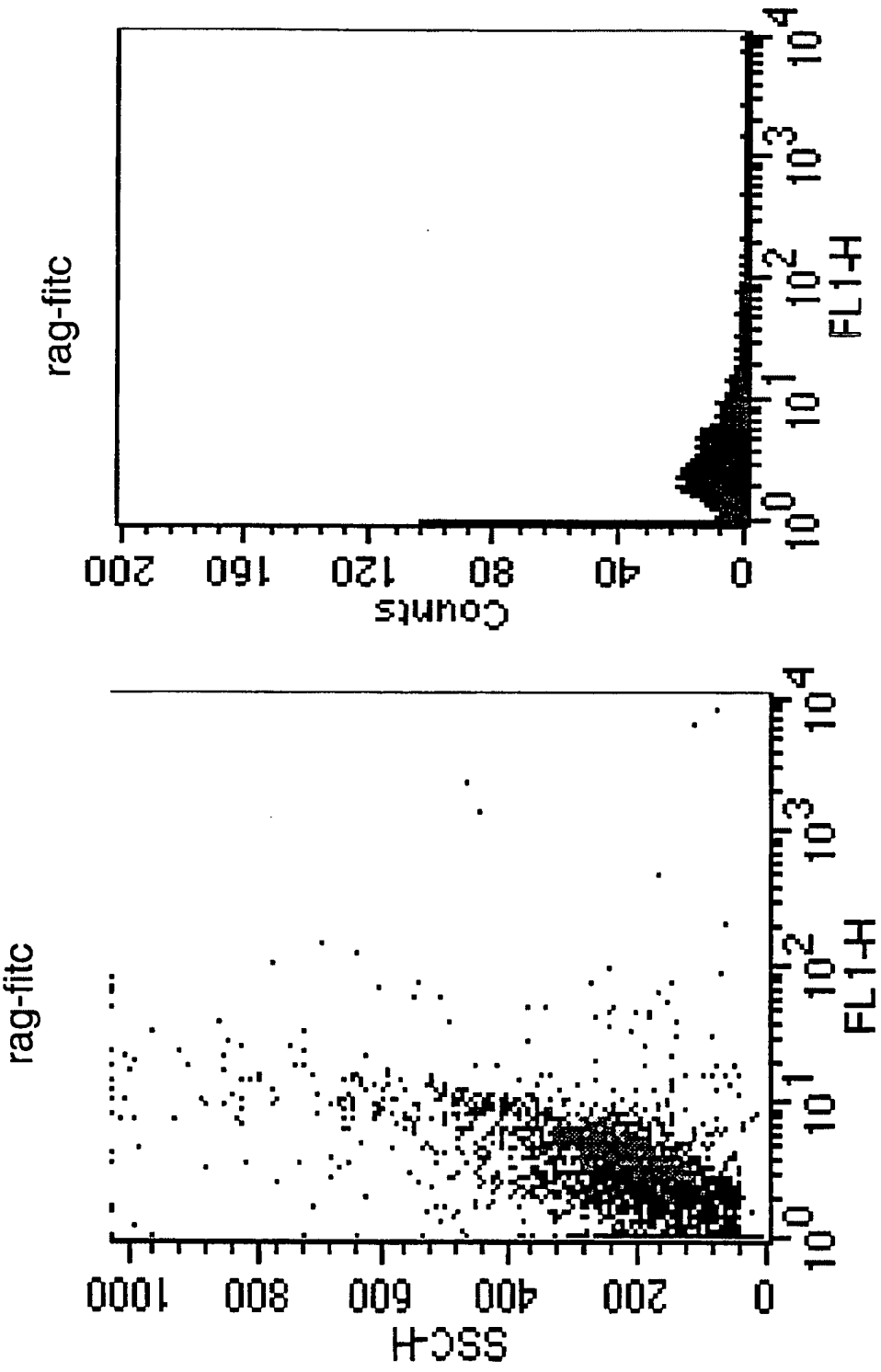


Figure 15B

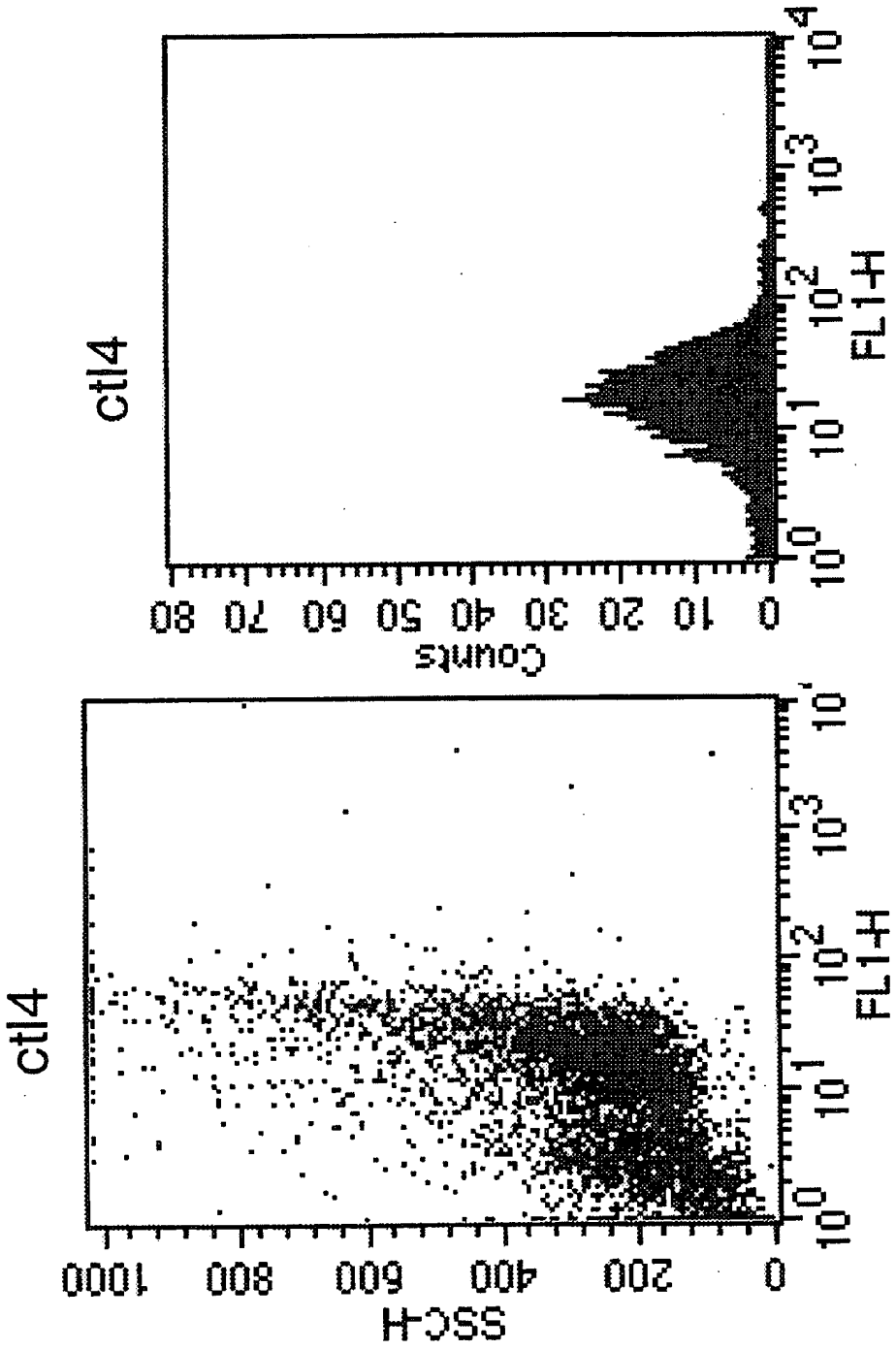

ctl4
FL1-H
Counts
0 10 20 30 40 50 60 70 80
ctl4
FL1-H
SSC-H
0 200 400 600 800 1000

Figure 15C

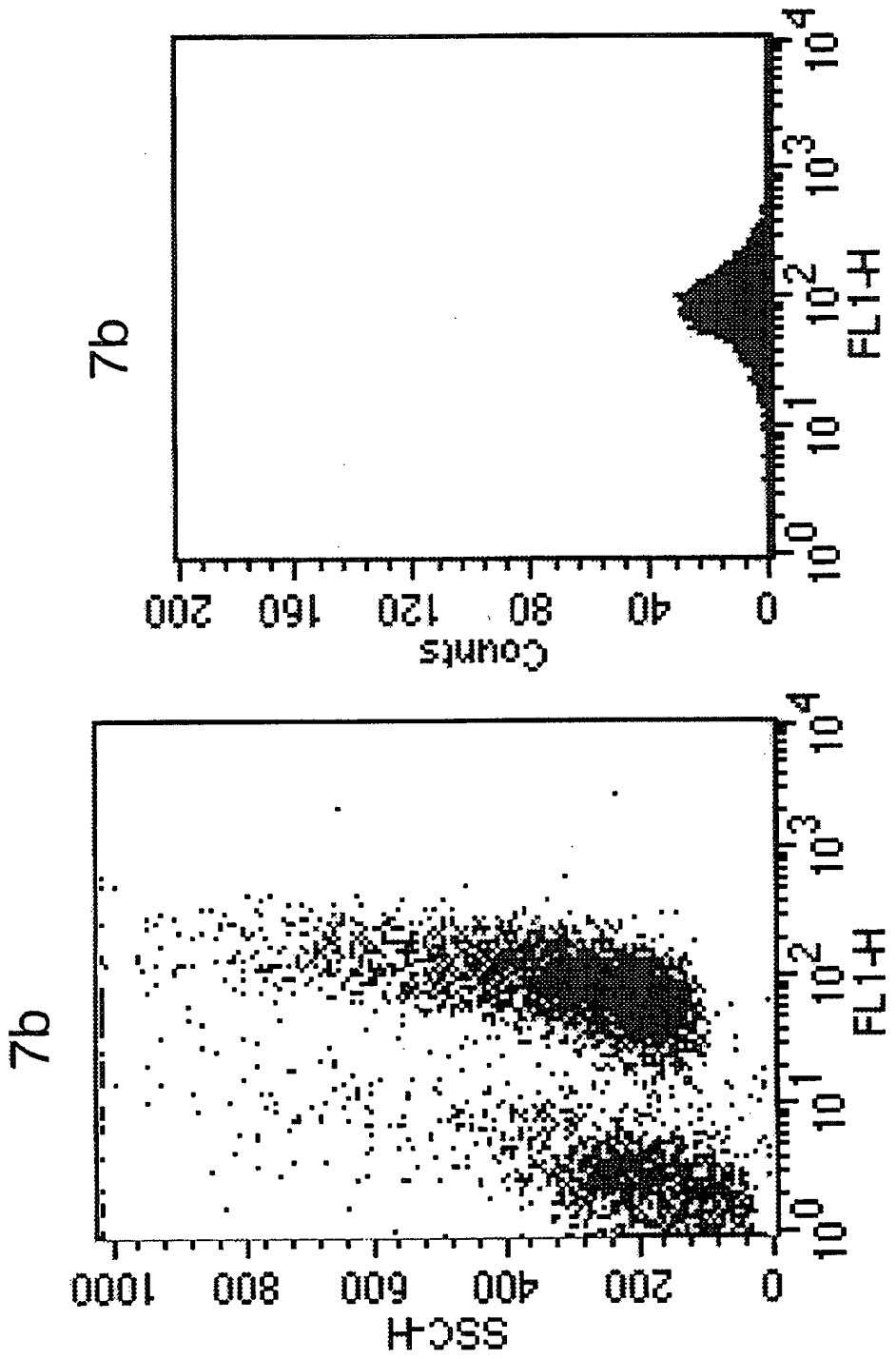

7b
Counts
0
40
80
120
160
200
FL1-H
10^0
10^1
10^2
10^3
10^4
7b
SSC-H
0
200
400
600
800
1000
FL1-H
10^0
10^1
10^2
10^3
10^4

Figure 15D

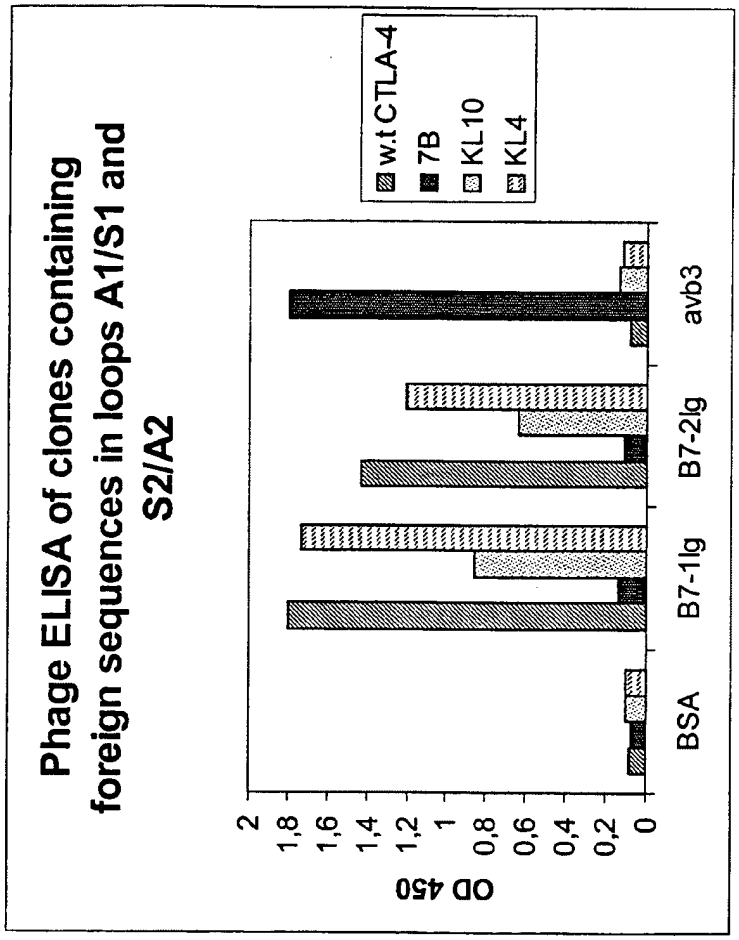

Phage ELISA of clones containing foreign sequences in loops A1/S1 and S2/A2
w.t CTLA-4
7B
KL10
KL4
OD 450
2
1,8
1,6
1,4
1,2
1
0,8
0,6
0,4
0,2
0
BSA
B7-1lg
B7-2lg
avb3

Figure 16

INTERNATIONAL SEARCH REPORT

Intern al Application No
PCT/EP 99/02283

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C07K14/705 C07K19/00 G01N33/68 C12N15/10

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C07K G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
E	EP 0 947 582 A (INNOGENETICS NV) 6 October 1999 (1999-10-06) abstract page 3, line 48-58 page 4, line 41-56 page 6, line 36 -page 8, line 12 page 96; claims ---	1-8,10, 11,15-22
E	WO 99 45110 A (DIATECH PTY LTD (AU); COIA G; GALANIS M; HUDSON PJ; IRVING RA; NUTTALL) 10 September 1999 (1999-09-10) page 2, line 16 -page 17, line 15; figure 6 page 29, line 5-26; example 9 page 36 -page 39; claims --- -/--	1-7,9, 10,15-22

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
7 January 2000	18/01/2000

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Macchia, G

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Inte. ›nal Application No
PCT/EP 99/02283

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 0 757 099 A (BRISTOL MYERS SQUIBB COMPANY (US)) 5 February 1997 (1997-02-05) abstract page 3, line 29 -page 9, line 38 page 24 -page 25; claims figures 1,3,6	1-3,5,6, 11, 15-18, 21,22
A	DARIAVACH P. ET AL.: "Human Ig superfamily CTLA-4 gene: chromosomal localization and identity of protein sequence between murine and human CTLA-4 cytoplasmic domains" EUROPEAN JOURNAL OF IMMUNOLOGY, vol. 18, December 1988 (1988-12), pages 1901-1905, XP000613323 page 1904; figure 5	3
T	NUTTALL S.D. ET AL.: "Design and expression of soluble CTLA-4 variable domain as a scaffold for the display of functional polypeptides" PROTEINS, vol. 36, no. 2, 1 August 1999 (1999-08-01), pages 217-227, XP000866035 the whole document	

2

INTERNATIONAL SEARCH REPORT

International application No.
PCT/EP 99/ 02283

B x I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Remark: Although claim 22
is directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

.formation on patent family members

Interr. ıal Application No
PCT/EP 99/02283

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0947582 A	06-10-1999	NONE	
WO 9945110 A	10-09-1999	NONE	
EP 0757099 A	05-02-1997	US 5773253 A	30-06-1998
		AU 696664 B	17-09-1998
		AU 6059096 A	30-01-1997
		CA 2181394 A	22-01-1997
		JP 9202800 A	05-08-1997
		NO 963018 A	22-01-1997